



Development and Implementation of Novel Genetic Tools for Investigation of FungalSecondary Metabolism

Holm, Dorte Koefoed

Publication date:
2013

Document Version
Publisher's PDF, also known as Version of record

[Link back to DTU Orbit](#)

Citation (APA):
Holm, D. K. (2013). *Development and Implementation of Novel Genetic Tools for Investigation of FungalSecondary Metabolism*. Technical University of Denmark.

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Development and Implementation of Novel Genetic Tools for Investigation of Fungal Secondary Metabolism

Dorte Marie Koefoed Holm

PhD thesis

Technical University of Denmark
Department of Systems Biology
Center for Microbial Biotechnology
Eukaryotic Molecular Cell Biology group

Preface

This study was carried out at the Center for Microbial Biotechnology, Department of Systems Biology, Technical University of Denmark, in the period 15 August 2010 to 14 August 2013. The study was supported by grant 09-064967 from the Danish Council for Independent Research, Technology, and Production Sciences.

There are many people I want to thank, and for many reasons. First and foremost my supervisors: Professor Uffe H. Mortensen, Professor Jens C. Frisvad, Assistant Professor Rasmus J. N. Frandsen and Senior Scientist Michael Lyng Nielsen (now with Novozymes A/S) who have all been there for me whenever I needed help, guidance or moral support. For that I am very grateful.

I want to thank Professor Stephen A. Osmani at the Department of Molecular Genetics, Ohio State University (OH, USA) for welcoming me into his lab, and I thank everyone in the Osmani Lab for making me part of your extended family and for being amazing co-workers and friends. Especially, I want to thank Aysha Osmani for helping me in every possible manner, inside and outside the lab.

Another thank you goes to Head of Department Hans van den Brink and PhD student Jesper Langholm Jensen at the Department of Enzymes and Innovation at Chr. Hansen A/S (Hørsholm, Denmark) and Kenneth S. Bruno at Pacific Northwest National Laboratory (Richland, WA, USA) for a good and honest collaboration on *A. niger* projects. Also, I thank Tim Hobley and Wickie Søndergaard (DTU Food) for an exciting collaboration on *M. gryphiswaldense*.

A special thank you goes to the colleagues and friends with whom I have collaborated closely: Diana C. Anyaogu, Lene M. Petersen, Paiman Khorsand-Jamal, Andreas Klitgaard, Thomas O. Larsen, Peter B. Knudsen, Jakob B. Nielsen, and Martin E. Kogle. I also thank Mikael R. Andersen for providing good advice, friendship, and gene expression data. I also owe a great thank you to all the students that I have supervised during my studies, emphasising Zofia D. Jarczynska who has been with me for a long time and done a terrific job!

I am very grateful to Kristian F. Nielsen for always critically engaging in new and exciting projects, for many interesting conversations, and for spicing up the social activities at the center. Thank you all at the center formerly known as CMB for making such a great working environment and for being my friends and extended family for the past 6 years.

Finally, I want to thank my non-CMB friends and family for supporting me throughout my studies - especially Andreas and Gilbert who have put up with me for better and for worse.

Abstract

Filamentous fungi are eukaryotic multicellular organisms, which are essential in our ecosystem for degradation of organic matter. Additionally, filamentous fungi produce an arsenal of small, bioactive molecules, *secondary metabolites*, which may be exploited for commercial purposes e.g. as pigments, fungicides, insecticides, pharmaceuticals etc. In this study I have examined the possibilities for elucidating the biosyntheses of fungal secondary metabolites, as well as identifying novel compounds, by application of genetic engineering techniques. I have focused on three levels of genetic engineering in three different filamentous fungi, all belonging to the genus *Aspergillus*: *A. nidulans*, *A. niger*, and *A. aculeatus*. *A. nidulans* is a well-characterized eukaryotic model system for which many genetic tools have been developed. *A. niger* is a well-characterized cell factory, however, as production strains have traditionally been constructed by random mutagenesis or directed evolution approaches only few genetic tools are publicly available. *A. aculeatus* is an industrial enzyme producer, which was only recently genome-sequenced. No genetic tools exist for this organism, and only very little is known about its potential to produce secondary metabolites.

In this study I have examined parts of the secondary metabolism in all three fungi, using different approaches. For *A. nidulans* we have constructed an entire genome-wide polyketide synthase (PKS) deletion library, which has linked the known compounds austinol and dehydroaustinol to the corresponding PKS and we have identified a new group of orsellinic acid-derived compounds. Also, I have successfully over-expressed individual *A. nidulans* transcription factors and phosphatases in an effort to perturb regulation of secondary metabolism and identify new compounds. Furthermore, *A. nidulans* was used as a host for examination of secondary metabolism by heterologous expression of secondary metabolite genes from *A. niger*, *A. aculeatus*, and *A. terreus*.

In *A. niger* I have investigated the products of all 37 putative PKSs by heterologous expression of the genes in *A. nidulans*. This approach identified the products of three PKSs, which were subsequently examined in detail in *A. niger*. One of the products was 6-methylsalicylic acid (6-MSA), which proved to be the precursor to the meroterpenoid yanuthone D, a compound that was previously hypothesized to be of mixed terpenoid and shikimic acid origin. Individual gene deletions in this gene cluster mapped the entire biosynthetic pathway towards yanuthone D, meanwhile identifying six novel yanuthones, which were not previously described.

The other two products both turned out to be the yellow pigment YWA1, which has previously been characterized from *A. nidulans* and *A. fumigatus*. YWA1 serves as the precursor to both naphtho- γ -pyrones and black melanin (DHN-melanin), which is the pigment that provides the characteristic black color of *A. niger*. Intriguingly, only one of these PKS genes (*napA*) is placed in a typical secondary metabolite gene cluster, but this PKS gene appears to be silenced.

For investigation of secondary metabolism in fungi with no available genetic tools, I have applied two different approaches: heterologous expression of secondary metabolite genes in *A. nidulans* and/or plasmid-based gene expression in the native fungus. Heterologous expression in *A. nidulans* was used for successful reconstruction of the (+)-geodin pathway, by transfer of the entire (+)-geodin cluster from *A. terreus*. For

A. aculeatus, I have used a combination of the two approaches: I heterologously expressed a putative 6-MSA synthase from *A. aculeatus* in *A. nidulans*, verifying that the gene indeed encoded a 6-MSA synthase. Next, a transcription factor embedded in the putative gene cluster, was over-expressed from a plasmid-based expression platform, which was developed in this study. Expression of the transcription factor triggered activation of a biosynthetic pathway, which appear to be involved in production of five novel compounds. Feeding experiments with fully ^{13}C labelled 6-MSA, produced in this study, confirmed that 6-MSA was incorporated into three of these compounds.

Based on these results, I conclude that the combination of different interdisciplinary approaches is a very powerful tool for investigating fungal secondary metabolism. The results of my studies also suggest that continuing optimization of existing tools as well as development of new tools is equally important as mere application of existing tools.

Sammenfatning

Filamentøse svampe er eukaryote multicellulære organismer der er essentielle for vores økosystem ved nedbrydning af dødt organisk materiale. Derudover er filamentøse svampe også i stand til at producere et arsenal af små, bioaktive molekyler, kaldet sekundære metabolitter, som kan udnyttes til kommercielle formål fx som pigmenter, svampegifte, insektgifte, lægemidler mv. Igennem dette studium har jeg undersøgt mulighederne for at belyse biosyntesen af sekundære metabolitter såvel som at identificere nye stoffer, ved at benytte forskellige genetiske værktøjer. Jeg har fokuseret på tre niveauer af genteknologi i tre forskellige filamentøse svampe som alle hører til genus *Aspergillus*: *A. nidulans*, *A. niger* og *A. aculeatus*. *A. nidulans* er et velkarakteriseret eukaryot modelsystem hvortil der er udviklet mange genetiske værktøjer. *A. niger* er en velkarakteriseret cellefabrik, men da produktionsstammer traditionelt set er lavet ved tilfældig mutagenese eller directed evolution er der kun få genetiske værktøjer offentligt til rådighed. *A. aculeatus* er en industriel enzymproducer som kun for nyligt er blevet genomsekventeret. Der eksisterer ingen genetiske værktøjer til denne organisme og kun meget lidt er kendt omkring dens potentiale til at producere sekundære metabolitter.

I dette studium har jeg undersøgt dele af den sekundære metabolisme i alle tre svampe, ved hjælp af forskellige fremgangsmåder. I *A. nidulans* har vi konstrueret et fuldgenom polyketidsyntase (PKS) gendelektionsbibliotek, som har koblet de kendte stoffer austinol og dehydroaustinol til deres respektive PKS og vi har identificeret en ny gruppe af orsellinsyre-derivede stoffer. Jeg har også med succes overudtrykt individuelle *A. nidulans* transkriptionsfaktorer og fosfataser i et forsøg på at perturbere reguleringen af den sekundære metabolisme for at finde nye stoffer. Derudover har jeg brugt *A. nidulans* som værtsorganisme i en undersøgelse af den sekundære metabolisme ved heterologt udtryk af sekundære metabolitgener fra *A. niger*, *A. aculeatus* og *A. terreus*.

I *A. niger* har jeg undersøgt produkterne fra samtlige 37 PKSer ved heterologt udtryk i *A. nidulans*. Ved hjælp af denne fremgangsmåde har jeg identificeret produkterne fra tre PKSer, som efterfølgende er undersøgt nærmere i *A. niger*. Et af produkterne var 6-methylsalicylsyre (6-MSA) som viste sig at være byggestenen for meroterpenoidet yanuthone D, et stof som tidligere var forudsagt til at stamme fra shikimisyre. Individuelle deletioner i dette genkluster har kortlagt hele biosyntesevejen mod yanuthone D, og samtidig har vi identificeret seks nye yanuthoner, som ikke er beskrevet tidligere.

De to andre produkter viste sig begge at være YWA1, som tidligere er karakteriseret fra *A. nidulans* og *A. fumigatus*. YWA1 er byggestenen for både naphtho- γ -pyroner og sort melanin (DHN-melanin) som er det sorte pigment der giver *A. niger* sin karakteristiske sorte farve. Genetiske analyser har vist at et af disse PKS gener, *napA*, er placeret i et typisk sekundær metabolitkluster men denne PKS ser ikke ud til at være udtrykt.

For at undersøge den sekundære metabolisme i svampe, hvor der ikke findes nogen genetiske værktøjer, har jeg brugt to forskellige fremgangsmåder: heterologt udtryk af sekundære metabolitgener i *A. nidulans* og/eller plasmidbaseret genudtryk i den native svamp. Heterologt udtryk i *A. nidulans* er med succes brugt til at rekonstruere hele (+)-geodin biosyntesevejen ved at overføre hele (+)-geodin genklusteret fra *A. terreus*.

I *A. aculeatus* har jeg anvendt en kombination af de to fremgangsmåder: jeg har udtrykt en formodet 6-MSA syntase fra *A. aculeatus* i *A. nidulans* og bekræftet at genet var en 6-MSA syntase. Derefter har jeg overudtrykt en transkriptionsfaktor fra genklusteret, via en plasmidbaseret platform der er blevet udviklet som en del af dette projekt. Udtrykket af transkriptionsfaktoren aktiverede en ukendt biosyntesevej som ser ud til at være ansvarlig for produktion af fem nye stoffer. Fødeforsøg med ^{13}C -mærket 6-MSA, der blev produceret som en del af dette projekt, bekræftede at 6-MSA inkorporeres i tre af de nye stoffer.

Baseret på disse resultater konkluderer jeg at kombinationen af interdisciplinære fremgangsmåder er et meget stærkt værktøj når man undersøger sekundær metabolisme. Resultaterne fra mine studier viser også at fortsat optimering af eksisterende værktøjer, samt udvikling af nye værktøjer, er mindst lige så vigtigt som kun at udnytte de værktøjer der allerede er til rådighed.

Publications

Peer-reviewed papers

- **Dorte K. Holm**, Lene M. Petersen, Andreas Klitgaard, Peter B. Knudsen, Zofia D. Jarczyska, Kristian F. Nielsen, Charlotte H. Gotfredsen, Thomas O. Larsen, and Uffe H. Mortensen: Molecular and chemical characterization of the biosynthesis of the 6-MSA derived meroterpenoid yanuthone D in *Aspergillus niger*. JACS, *submitted for publication* (Chapter 10).
- Morten T. Nielsen, Jakob B. Nielsen, Diana C. Anyaogu, **Dorte K. Holm**, Kristian F. Nielsen., Thomas O. Larsen and Uffe H. Mortensen: Heterologous Reconstitution of the Intact Geodin Gene Cluster in *Aspergillus nidulans* Through a Simple and Versatile PCR Based Approach. PLoS ONE **2013**, 8, 8:e72871 (Chapter 11).
- Marie Louise Klejnstrup, Rasmus J.N. Frandsen, **Dorte K. Holm**, Morten T. Nielsen, Uffe H. Mortensen, Thomas O. Larsen, Jakob B. Nielsen: Review: Genetics of polyketide metabolism in *Aspergillus nidulans*, Metabolites **2012**, 2, 100-133 (Chapter 5).
- Michael L. Nielsen, Jakob B. Nielsen, Christian Rank, Marie L. Klejnstrup, **Dorte K. Holm**, Katrine H. Brogaard, Bjarne G. Hansen, Jens C. Frisvad, Thomas O. Larsen and Uffe H. Mortensen: Genome wide polyketide synthase gene deletion library in *Aspergillus nidulans*, FEMS Microbiol Lett **2011**, 1-10 (Chapter 4).
- Lars R. Olsen, Niels B. Hansen, Mads T. Bonde, Hans J. Genée, **Dorte K. Holm**, Simon Carlsen, Bjarne G. Hansen, Kiran R. Patil, Uffe H. Mortensen, Rasmus Wernersson: PHUSER (Primer Help for USER): A novel tool for USER fusion primer design, Nucleic Acids Research **2011**, 1-7; doi: 10.1093/nar/gkr394 (Chapter 7).

Manuscripts in preparation

- Lene M. Petersen*, **Dorte K. Holm***, Charlotte H. Gotfredsen, Thomas O. Larsen, and Uffe H. Mortensen: Over-expression of Transcription Factor in *Aspergillus aculeatus* Activates Biosynthetic Pathway and Reveals Novel Compounds Acurin A and B. *Joint first authors (Chapter 12).
- Jakob B. Nielsen, Marie L. Klejnstrup, Paiman Khorsand-Jamal, Anna Kabat, **Dorte K. Holm**, Michael L. Nielsen, Jens C., Frisvad, Charlotte H. Gotfredsen, Thomas O. Larsen, and Uffe H. Mortensen: The Pathway Leading to Nitrogen Incorporation into 3-MethylOrsellinic Acid Derived Natural Products in *Aspergillus nidulans*.
- **Dorte K. Holm**, Rasmus J.N. Frandsen, Kristian F. Nielsen, and Uffe H. Mortensen: Identification of the naphtho- γ -pyrone gene cluster in *Aspergillus niger*. Intended for submission to Chemistry and Biology (Chapter 9).

Conference contributions

- **Dorte K. Holm**, Lene M. Petersen, Andreas Klitgaard, Zofia Jarczynska, Thomas O. Larsen, and Uffe H. Mortensen: Elucidation of the Biosynthesis of Yanuthone D in *Aspergillus niger*. International Congress on Natural Products Research (ICNPR), New York, NY, USA (2012), abstract published in *Planta Medica* — 2012, Volume 78, P1069.
- Jakob B. Nielsen, Marie L. Klejnstrup, Paiman Khorsand-Jamal, **Dorte K. Holm**, Michael L. Nielsen, Anna M. Kabat, Christian Rank, Charlotte H. Gotfredsen, Thomas O. Larsen, and Uffe H. Mortensen: Characterization of the AN6448 cluster in *Aspergillus nidulans*. International Congress on Natural Products Research (ICNPR), New York, NY, USA (2012), abstract published in *Planta Medica* — 2012, Volume 78, P1113.
- Andreas Klitgaard, **Dorte K. Holm**, Jens Chr. Frisvad, and Kristian Fog Nielsen: Screening of *Aspergillus nidulans* metabolites from habitat mimicking media using LC-DAD-TOFMS system. International Congress on Natural Products Research (ICNPR), New York, NY, USA (2012), abstract published in *Planta Medica* — 2012, Volume 78, P167.
- Lene M. Petersen, **Dorte K. Holm**, Charlotte Held Gotfredsen, Uffe H. Mortensen, and Thomas O. Larsen: Discovery of novel secondary metabolites in *Aspergillus aculeatus*. International Congress on Natural Products Research (ICNPR), New York, NY, USA (2012), abstract published in *Planta Medica* — 2012, Volume 78, P1159-1160.
- **Dorte K. Holm**, Lene M. Petersen, Thomas O. Larsen, and Uffe H. Mortensen: Activation of a silent meroterpenoid pathway in *Aspergillus aculeatus*. 7th Danish Conference on Biotechnology and Molecular Biology, Vejle, Denmark (2012).
- Thomas O. Larsen, Marie L. Klejnstrup, Jakob B. Nielsen, **Dorte K. Holm**, Lene M. Petersen, Andreas Klitgaard, Kristian F. Nielsen, Mikael R. Andersen, and Uffe H. Mortensen: Chemical analysis of a genome wide polyketide synthase gene deletion library in *Aspergillus nidulans*. Directing Biosynthesis III, Nottingham, UK (2012).
- **Dorte K. Holm**, Bjarne G. Hansen, Thomas O. Larsen, and Uffe H. Mortensen: Expression of polyketide and non-ribosomal peptide synthases in *A. nidulans*. 6th Danish Conference on Biotechnology and Molecular Biology, Vejle, Denmark (2011).
- **Dorte K. Holm**, Bjarne G. Hansen, Thomas O. Larsen, and Uffe H. Mortensen: Heterologous Expression of Polyketide Synthases in *Emericella nidulans*. 26th Fungal Genetics Conference, Asilomar, CA, USA (2011).
- **Dorte K. Holm**, Bjarne G. Hansen, Thomas O. Larsen, and Uffe H. Mortensen: Heterologous Expression of Polyketide Synthases in *Emericella nidulans*. 8th Asperfest Conference, CA, USA (2011).

Abbreviations

ABC transporter	ATP-binding cassette transporter
ACP	acyl carrier protein
AMA1	Autonomous Aaintenance in <i>Aspergillus</i>
AspGD	Aspergillus Genome Database
AT	acyl transferase
BLAST	Basic Local Alignment Search Tool
bp	base pairs
BSA	bovine serum albumin
cDNA	complementary DNA
CDS	coding sequence
CoA	Coenzyme A
CYA	Czapek Yeast Agar
DAD	Diode Array Detector
DGLE	DNA glycosylase-lyase Endonuclease VIII
DH	dehydrogenase/dehydratase
DOX	doxycycline hyclate/doxycycline hydrochloride hemiethanolate hemihydrate
DNA	deoxyribonucleic acid
DTU	Technical University of Denmark
ESI-	Electrospray Ionization, negative mode
ESI+	Electrospray Ionization, positive mode
ER	enoyl reductase
5-FOA	5-fluoroorotic acid
gDNA	genomic DNA
GFP	green fluorescent protein
HPLC	high performance liquid chromatography
HR	homologous recombination
HRMS	high-resolution mass spectrometry
HTP	high throughput
IS	integration site
IS1	Insertion Site 1 [Hansen et al., 2011]
JGI	DOE Joint Genome Institute

Continued on next page

Abbreviations – Continued

kb	kilobase pairs
KR	Keto reductase
KS	<i>beta</i> -ketoacyl synthase
LC	liquid chromatography
MEA	Malt Extract Agar
MM	minimal medium
MPA	metabolite profile analysis
mRNA	messenger RNA
MS	mass spectrometry
MS/MS	tandem mass spectrometry
MT	methyl transferase
m/z	mass-to-charge ratio
6-MSA	6-methylsalicylic acid
MQ	ultrafiltrated water
NEB	New England Biolabs
NHEJ	non-homologous end-joining
NMR	nuclear magnetic resonance
NRP	non-ribosomal peptide
NRPS	non-ribosomal peptide synthase
NR-PKS	non-reducing PKS
nt	nucleotide(s)
ORF	open reading frame
<i>P_{alcA}</i>	promoter region of <i>alcA</i> , 0.8 kb
<i>P_{gpdA}</i>	promoter region of <i>gpdA</i> , 2.3 kb
<i>P_{gpdA}_{0.5kb}</i>	promoter region of <i>gpdA</i> , 0.5 kb
<i>P_{O7}</i>	O7 promoter, inducible with DOX
PCR	Polymerase Chain Reaction
PK	polyketide
PKS	polyketide synthase
PNNL	Pacific Northwest National Laboratories
ppm	parts per million
PPT	4'-phosphopanthetheinyl
qRT-PCR	quantitative real-time PCR

Continued on next page

Abbreviations – Continued

RFP	red fluorescent protein
RNA	ribonucleic acid
RNAi	interfering ribonucleic acid
RT	retention time
SAM	S-adenosylmethionine
SM	secondary metabolite
TE	thioesterase
TF	transcription factor
TOF	time-of-flight
TOFMS	time-of-flight mass spectrometry
TR	thioester reductase
TS	targeting sequence
UDG	Uracil DNA glycosylase
UE	uracil excision
UEC	uracil excision cassette
UHPLC	ultra high performance liquid chromatography
UTR	untranslated region
UV	ultraviolet
UV/VIS	ultraviolet/Visual
X-gal	5-bromo-4-chloro-3-indoyl- β -D-galactopyranoside
YES	yeast extract sucrose medium
YFG	your favourite gene

List of Figures

1.1	Overview of the connections between method development and implementation thereof	6
1.2	Filamentous growth of <i>Aspergilli</i>	8
1.3	Asexual and parasexual life cycle of <i>Aspergilli</i>	9
1.4	<i>Aspergilli</i> used in this study	10
1.5	The toolbox for genetic engineering	12
1.6	The pathway towards green melanin in <i>A. nidulans</i>	13
1.7	Generation of single-stranded overhangs by uracil base excision	17
1.8	Integration into IS1	18
1.9	Principle of UE cloning	19
1.10	Classes of fungal secondary metabolites	21
1.11	Types of polyketide synthases	22
1.12	Schematic illustration of synthesis by a fungal iterative PKS	25
1.13	Reduction of the ketide unit	26
1.14	Schematic of a typical gene cluster	27
1.15	Approaches for generating intron-free DNA	30
1.16	Principle of the UHPLC-DAD-TOFMS system	31
1.17	Base peak chromatogram (BPC) of <i>A. niger</i> extract	32
1.18	UV spectra of pyranonigrin A, aurasperone B, and yanuthone D	33
1.19	Mass spectra of pyranonigrin A, aurasperone B, and yanuthone D	33
2.1	Transformation plate using the background-free integration system for <i>A. nidulans</i>	38
2.2	Production of YWA1 from six different integration sites in <i>A. nidulans</i>	39
3.1	Qualitative test of new promoters on X-gal plate	48
3.2	Test of <i>A. nidulans</i> integration sites using <i>lacZ</i> reporter gene	52
3.3	Test of <i>A. nidulans</i> integration sites using <i>wA</i> reporter gene	52
3.4	Test of <i>A. niger</i> integration site 2	54
3.5	Test of <i>A. niger</i> integration sites 3 and 4	55
3.6	AMA1-based expression vectors developed in this study	56

3.7	Test of AMA1 expression vectors using a color screen	57
3.8	CMBU1111 and the background-free expression vector pDH56	58
3.9	Background-free transformation system in <i>A. nidulans</i>	59
6.1	Impact of growth medium composition on the <i>A. nidulans</i> polyketome	116
6.2	Background-free cloning vector pDH124 employing TetON/OFF	117
6.3	Macromorphologies of phosphatase over-expression strains	118
6.4	Histogram of colony diameters of phosphatase over-expression strains	119
6.5	Macromorphologies of TF over-expression strains	120
6.6	BPC (ESI+) chromatogram of six selected strains cultivated on 1.0 $\mu\text{g}/\text{ml}$ DOX	123
6.7	BPC (ESI-) chromatogram of six selected strains cultivated on 1.0 $\mu\text{g}/\text{ml}$ DOX	124
6.8	DAD chromatogram of six selected strains cultivated on 1.0 $\mu\text{g}/\text{ml}$ DOX	125
6.9	Selection of known <i>A. nidulans</i> compounds	126
6.10	UV spectra of four compounds which were not identified by dereplication	127
6.11	Structure of the recently reported asperserin A	128
8.1	Selected <i>A. niger</i> secondary metabolites	151
8.2	Cloning vector pCMBU1111	152
8.3	Phenotypes of <i>A. nidulans</i> expressing the ASPNIDRAFT_191422 and ASPNIDRAFT_56896 genes	152
8.4	Dereplication of new compounds	154
8.5	BPC and EIC chromatograms of the 44005 Δ and the NRPS Δ strains	155
8.6	Proposed biosynthesis of asperrubrol	156
8.7	EIC chromatograms of IS1-50432 and reference compared to a cinnamic acid standard	157
H.1	Distribution of secondary metabolite synthases in <i>A. aculeatus</i>	390
H.2	Gene prediction for putative PKS 46489	394
H.3	Annotation of putative PKS 46489	394
H.4	Gene prediction for putative PKS 1882131	395
H.5	Annotation of putative PKS 1882131	396
H.6	Activity of the omega-hydroxypalmitate O-feruloyl transferase (OHFT)	397
H.7	Phylogenetic tree of <i>A. aculeatus</i> PKSs	398

List of Tables

1.1	Assessment of current methods for substrate generation	16
1.2	Overview of domains in the iterative type I PKS.	24
3.1	<i>A. nidulans</i> promoters analyzed in this study	49
3.2	Insertion sites for <i>A. nidulans</i>	51
3.3	Insertion sites for <i>A. nidulans</i>	53
3.4	Fungal strains used and constructed in this study	61
3.5	Primers used in this study	64
3.6	Plasmids used in this study	67
6.1	Colony diameters of phosphatase over-expression strains	119
6.2	Colony diameters of TF over-expression strains	121
6.3	Fungal strains used and constructed in this study	130
6.4	Primers used in this study	133
6.5	Plasmids used and constructed in this study	138
8.1	<i>A. niger</i> PKSs	157
8.2	<i>A. nidulans</i> strains used and constructed in this study	161
8.3	Primers used in this study	163
8.4	Plasmids used and constructed in this study	167
H.1	Secondary metabolite synthases in <i>A. aculeatus</i>	391

Contents

1	Introduction	5
1.1	General introduction to thesis	5
1.2	Filamentous fungi	7
1.2.1	Life cycle of Aspergilli	7
1.3	Genetic engineering	12
1.3.1	Tools for genetic engineering in <i>Aspergillus nidulans</i>	12
1.3.2	Tools for genetic engineering in <i>Aspergillus niger</i>	14
1.4	Uracil excision cloning	14
1.4.1	IS1 expression platform	18
1.5	Secondary metabolism in <i>Aspergillus</i>	20
1.5.1	Fungal polyketides	21
1.5.2	Synthesis by the type I iterative polyketide synthases	25
1.5.3	Secondary metabolite gene clusters	26
1.5.4	Transcription factors	26
1.5.5	Heterologous expression of secondary metabolite genes in <i>Aspergillus</i>	28
1.6	Metabolite profile analysis	30
1.6.1	Liquid chromatography	31
1.6.2	Diode Array Detection (DAD)	32
1.6.3	Mass Spectrometry (MS)	33
2	Overall results and discussion	35
2.1	Investigation of secondary metabolism in <i>A. nidulans</i> : a fungus with a large genetic toolbox	35
2.1.1	Systematic over-expression of phosphatases and TF triggers secondary metabolite production in <i>A. nidulans</i>	36
2.1.2	Assessment of novel genetic tools	37
2.2	Investigation of secondary metabolism in <i>A. niger</i> : a fungus with few genetic tools	39
2.2.1	Heterologous expression links three PKSs to compounds	40
2.2.2	Combining several approaches reveals the biosynthesis of yanuthone D in <i>A. niger</i>	41

2.3	Investigation of secondary metabolism in fungi with no available genetic tools	42
2.3.1	<i>A. nidulans</i> as a host for heterologous gene expression	44
2.4	Concluding remarks and future perspectives	45
3	Development of novel tools for genetic engineering in <i>Aspergilli</i>	47
3.1	Results and discussion	47
3.1.1	Expanding the library of promoters for <i>A. nidulans</i>	47
3.1.2	Development of five new integration sites in <i>Aspergillus nidulans</i>	50
3.1.3	Development of four integration sites in <i>Aspergillus niger</i>	53
3.1.4	Development of a non-integrative (AMA1-based) expression platform for <i>Aspergillus</i> Spp.	55
3.1.5	Development of a HTP gene expression platform for <i>A. nidulans</i>	57
3.2	Concluding remarks	60
3.3	Experimental	61
3.3.1	Fungal strains	61
3.3.2	Media	62
3.3.3	PCR, cloning and transformation	62
4	A genome-wide polyketide synthase deletion library uncovers novel genetic links to polyketides and meroterpenoids in <i>Aspergillus nidulans</i>	69
5	Genetics of Polyketide Metabolism in <i>Aspergillus nidulans</i>	80
6	Identification of novel secondary metabolites by over-expression of phosphatases and transcription factors from <i>A. nidulans</i>	115
6.1	Introduction	115
6.2	Results and discussion	117
6.2.1	Over-expression of phosphatases affects growth rate and morphology	117
6.2.2	Over-expression of TFs results in several new phenotypes	119
6.2.3	Metabolite profile analysis reveals production of novel metabolites	122
6.3	Concluding remarks	129
6.4	Experimental procedures	130
6.4.1	Fungal strains	130
6.4.2	Media	132
6.4.3	PCR, cloning and transformation	132
6.5	Chemical analysis of strains	140
7	PHUSER (Primer Help for USER)	141

8	Heterologous expression of polyketide synthases in <i>A. nidulans</i>	149
8.1	Introduction	149
8.2	Results and discussion	150
8.2.1	Heterologous expression of 37 putative PKSs links three PKSs to a compound	150
8.2.2	Co-expression with discrete thioesterases did not provide any new compounds	153
8.2.3	Gene deletion links 44005 to asperrubrol	153
8.3	Assessment of <i>A. nidulans</i> as a host for heterologous PKS expression	159
8.3.1	Only three compounds were linked to a PKS	159
8.3.2	There are obvious bottlenecks in HTP strain construction	160
8.4	Experimental	161
8.4.1	Fungal strains	161
8.4.2	Media	162
8.4.3	PCR, cloning and transformation	162
8.4.4	Chemical analysis of strains	168
9	Identification of the naphtho-γ-pyrone gene cluster in <i>Aspergillus niger</i>	170
10	Molecular and chemical characterization of the biosynthesis of the 6-MSA derived meroterpenoid yanuthone D in <i>Aspergillus niger</i>	193
11	Heterologous Reconstitution of the Intact Geodin Gene Cluster in <i>Aspergillus nidulans</i> Through a Simple and Versatile PCR Based Approach	226
12	Over-expression of Transcription Factor in <i>Aspergillus aculeatus</i> Activates Biosynthetic Pathway and Reveals Novel Compounds Acurin A and B	237
A	Supporting information for Chapter 4	270
B	Supplementaries for Chapter 6	281
B.1	Large pictures of phosphatase and TF strains	281
B.2	Chromatograms of phosphatase and TF strains	291
C	Supporting information for Chapter 7	334
D	Supporting information for Chapter 9	339
E	Supporting information for Chapter 10	343
F	Supporting information for Chapter 11	373
G	Supporting information for Chapter 12	384

H	Investigation of the secondary metabolite potential in <i>Aspergillus aculeatus</i>	390
H.1	Identification of synthases	390
H.2	Annotation and assessment of polyketide synthases	393
H.2.1	46489 is a non-ordinary or non-functional PKS	393
H.2.2	120020 may be a type III PKS	395
H.2.3	1882131 is a non-ordinary PKS that contains a novel SAT domain	395
H.3	Phylogeny of the <i>A. aculeatus</i> PKSs	397

Chapter 1

Introduction

1.1 General introduction to thesis

Within filamentous fungi the genus *Aspergillus* is one of the most important in relation to humans: *Aspergillus fumigatus* is a human pathogen, *A. flavus* produces the carcinogenic aflatoxin, *A. oryzae* is widely used in Asia for production of traditional foods, *A. niger* is an industrial workhorse used for production of organic acids and enzymes, and *A. nidulans* is used as a model organism for eukaryotes.

As Aspergilli have such an impact on our lives, is it important that we possess the ability to study these organisms, whether it be in the context of finding a cure for fatal *Aspergillus* infections (aspergillosis) or in the context of producing compounds that are beneficial to humans e.g. enzymes, pigments or pharmaceuticals. During this study I have assessed the current available tools for performing genetic engineering in two very different Aspergilli: the model organism *A. nidulans* and the industrially important *A. niger* and tried to optimize and expand the tools and methods in relation to implementation of a high-throughput (HTP) setting for genetic engineering. I have successfully combined the development of new tools with implementation thereof, in order to investigate secondary metabolism, primarily in the three Aspergilli; *A. nidulans*, *A. niger* and *A. aculeatus*.

Figure 1.1 shows the connections between development of new tools, and the implementation thereof for investigation of secondary metabolism.

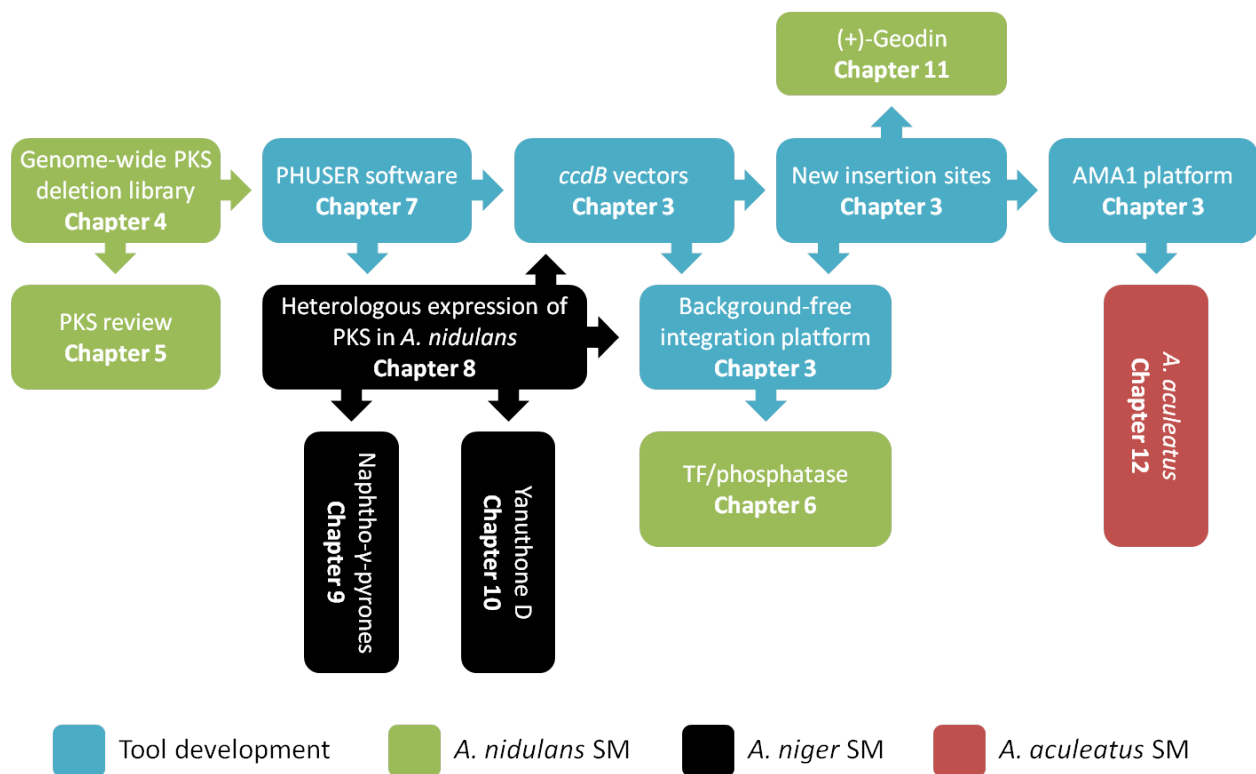


Figure 1.1: Overview of the connections between method development and implementation thereof. SM = secondary metabolism.

In this thesis, I first provide a short introduction to filamentous fungi, currently available genetic tools, and secondary metabolism (Chapter 1). This introduction will provide a background for reading and assessing all of the following chapters.

Following the introduction, I will very briefly present the overall results of this study and discuss them in the context of genetic approaches and fungal secondary metabolism in general (Chapter 2).

Subsequently, I briefly present the novel genetic tools that I have developed in order to expand and improve the current toolbox for genetic engineering in *A. nidulans* and *A. niger*, providing a basis for conducting both HTP fungal strain construction and in-depth investigations of cell biology (Chapter 3). The tools presented in this chapter have been applied for investigation of secondary metabolism throughout the study.

Then I first present the results of *A. nidulans* studies (Chapters 4, 5, and 6). Chapter 4 describes construction of a genome-wide polyketide synthase (PKS) deletion library for linking PKS genes to metabolites, and Chapter 5 comprises a review that summarizes the current state of linking PKS genes to metabolites in *A. nidulans* as per December 2012. Chapter 6 describes a HTP method for triggering metabolite production in *A. nidulans*, combining novel genetic techniques for over-expression of *A. nidulans* phosphatases and transcription factors.

Chapters 7, 8, 9, and 10 contain studies on *A. niger*: Chapter 7 contains a paper in which we describe the development of a web-based computer program, PHUSER, which enables the non-experienced user to

flexibly design PCR primers that are compatible with the uracil-excision (USERTM) cloning system that has been applied throughout this study. Chapter 8 describes heterologous expression of all PKS genes from *A. niger* in *A. nidulans*, linking three polyketide compounds to their PKSs. Based on these results, Chapters 9 and 10 describe in-depth studies on naphtho- γ -pyrones and yanuthone D biosyntheses, in *A. niger*.

Lastly, I present reconstruction of the (+)-geodin pathway of *A. terreus* in *A. nidulans* (Chapter 11) and activation of a biosynthetic pathway in *A. aculeatus* that resulted in identification of five novel compounds (Chapter 12). For each chapter, I have included a short introduction to the specific study, and for chapters comprising manuscripts in preparation, I have also included a description of work that needs to be finished before submitting the paper for publication. To avoid redundancy, a combined reference list for all chapters is provided in the end, except for chapters that contain published manuscripts.

1.2 Filamentous fungi

Filamentous fungi, or *mold fungi*, constitute a huge group of multicellular eukaryotes that are essential for our ecosystem, as they have a large role in biodegradation of organic matter. More than 100,000 fungal species are known, and it is estimated that as much as 1,400,000 species exist [Petersen, 2003]. Because of their natural ability to degrade organic material fungi are most commonly thought of as food spoilers, but they also have several other qualities that make them interesting organisms to study. One of the characteristics that make fungi stand out from other organisms is their ability to grow in very diverse environments: they are very tolerant to changes in pH, they degrade and consume a wide range of substrates, and they can survive in very harsh environments. Furthermore, many species possess the ability to form specialized structures, which can survive in a dormant state for many years until environmental conditions are appropriate for propagation.

Importantly, filamentous fungi are capable of producing an arsenal of bioactive small organic compounds. The possibility of exploiting this ability to improve human life was indeed realized when Alexander Fleming discovered that the filamentous fungus *Penicillium notatum* produced the antibacterial compound *penicillin* [Fleming, 1942, 1943], which became one of the most important medicines of the 20th century. Since then, many more compounds have been isolated from filamentous fungi, which possess bioactivities that can be exploited to human advantage e.g. the cholesterol-lowering statins and the antimicrobial and immunosuppressive mycophenolic acid. In fact, 64 % of all approved drugs are natural products or natural product mimics [Newman and Cragg, 2012].

1.2.1 Life cycle of Aspergilli

Filamentous fungi are multicellular organisms that form long tubular structures (filaments) called *hyphae* (see Figure 1.2). In Aspergilli, the hypha is divided into compartments by *septa*, which are perforated walls that allow for diffusion of nutrients, proteins, and even organelles. Thus the whole network of hyphae, the *mycelium*, is considered one single organism.

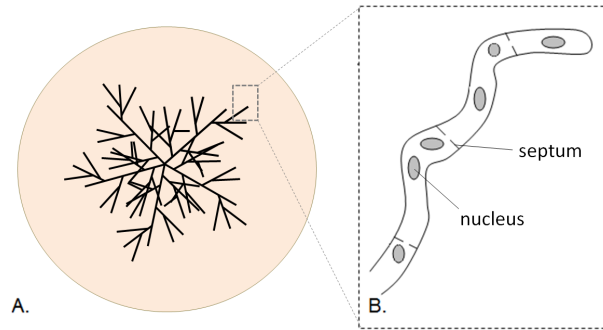


Figure 1.2: Filamentous growth of Aspergilli. The *mycelium* (A.) comprises a network of tubular filaments called *hyphae* (B.). In Aspergilli, hyphae contain perforated *septa* which form compartments within the hypha.

Aspergilli are able to propagate via different lifecycles: asexually, parasexually, and sexually (see Figure 1.3). However, not all Aspergilli have a known sexual cycle. The asexual life cycle begins with a single asexual spore called a *conidium*. If nutrients are available the conidium germinates to form the tadpole-like *germling* which matures to form a hypha and eventually the entire mycelium. From the hypha, a stalk will arise and form a structure called a *conidiophore* (see Figure 1.3). The conidiophore consists of a vesicle, coated with *metulae* and *phialides* which are specialized cells for production of conidia. The conidia will bud from the phialides in long chains, and the extremely hydrophobic conidia will be dispersed into the surrounding environment by wind, water, or by passing animals. Each of those conidia are then able to form a new mycelium if nutrition is available.

The parasexual life cycle allows the fungus to change ploidity and exchange nuclei. This means that different fungi of the same species can fuse their mycelia by a process called *anastomosis*, and propagate as one organism. In this case the "new" fungus, the *heterokaryon*, would contain two different and independent nuclei. In some cases, nuclei of both heterokaryons and homokaryons will fuse to form a diploid. Diploid fungi can propagate via the asexual cycle on equal terms as haploid fungi. Similarly, homokaryons, heterokaryons, and diploids can enter the sexual cycle, which will not be described in the present thesis.

Knowledge of the fungal life cycle allows researchers to make use of the different parts of the life cycle to their advantage. Asexual propagation is desired for basic cultivation of fungi, whether it be in a Petri dish or in a fermentor. For industrial purposes it is desirable to maintain asexual propagation for production of a given compound or protein, whereas it may be desirable for the molecular biologist to exploit meiosis in order to cross different strains and make hybrids. However, the presence of different life cycles can also be a disadvantage when handling fungi in the laboratory, as one has to take care that a given strain has not fused with other strains to form heterokaryons or changed ploidity.

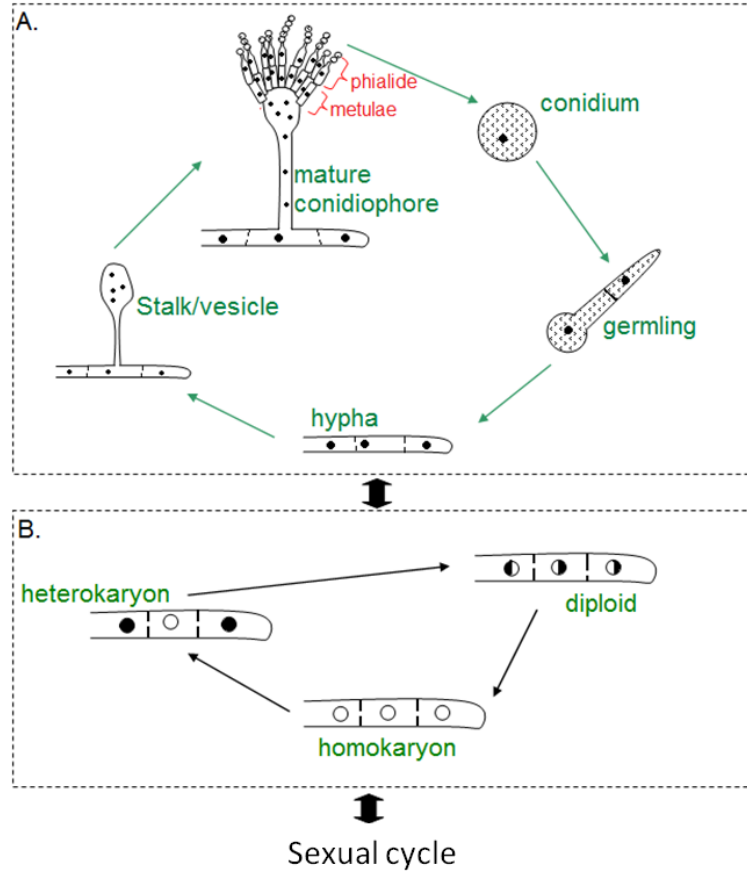


Figure 1.3: Asexual and parasexual life cycle of Aspergilli. **The asexual life cycle (A.)** begins with a single conidium which germinates to form the hypha. From the hypha a conidiophore, comprising a vesicle with metulae and phialides, will arise and conidia will bud from the conidiogenous phialides and form long chains of hydrophobic conidia. **The parasexual life cycle (B.)** allows the fungus to interchange between haploid and diploid stages, as well as existing as a heterokaryon: a haploid fungus containing two (or more) different, but intact, nuclei.

In this study I have chosen to investigate three very different fungi, which all belong to the Aspergilli: *A. nidulans*, which is a well-studied model organism for filamentous fungi, *A. niger*, which is an important industrial work-horse, and *A. aculeatus*, which is an industrial enzyme producer, which was only recently genome-sequenced by the DOE Joint Genome Institute (Figure 1.4).



Figure 1.4: Aspergilli used in this study.

A. nidulans is a filamentous fungus, which is considered an important model organism for filamentous fungi. *A. nidulans* was genetically characterized by Guido Pontecorvo in the 1950s, and due to his pioneering work within classic genetics *A. nidulans* is now a well-studied model organism for eukaryotes [Pontecorvo et al., 1953, Morris and Enos, 1992]. As a consequence of this, several tools for genetic engineering in *A. nidulans* exist, and for this reason this fungus is often the first choice for development of new methods and techniques for filamentous fungi. The details of the tools are described in Section 1.3.1.

As many tools exist for genetic engineering in *A. nidulans*, the *polyketome* (complete set of polyketides) was explored by different approaches: genome-wide PKS deletion, over-expression of genes, and site-directed mutagenesis. Furthermore, *A. nidulans* was used for heterologous expression of genes from organisms where little or no genetic tools exist(ed).

A. niger is a very important fungus for production of organic acids and enzymes. For that reason, *A. niger* has been studied extensively with respect to behavior in liquid cultures. Traditionally, fermentation organisms have been optimized for better morphology and productivity via random mutagenesis and random gene insertions, respectively, and for that reason not many genetic techniques are (publicly) available. Moreover, *A. niger* does not have a known sexual cycle, which makes it difficult to introduce or remove genetic alterations by sexual crossing, as it is possible in *A. nidulans*.

A. niger is known to produce a wide array of secondary metabolites, including the harmful fumonisins [Frisvad et al., 2007]. As only few genetic tools are available for this organism, I have explored the polyketome of *A. niger* using a combination of heterologous expression in *A. nidulans* and gene deletions in *A. niger*.

A. aculeatus is a recently genome-sequenced fungus, belonging to the group of black Aspergilli. *A. aculeatus* is a natural plant pathogen why it is able to produce enzymes for degradation of plant material e.g. cellulases and hemicellulases. As these enzymes are of great value for production of second generation bioethanol this fungus is used in industrial fermentations. Interestingly, only little research has been conducted to elucidate its potential of producing secondary metabolites. Exploring the secondary metabolism is possible by traditional mycological tools; cultivation on different media followed by metabolite profile analysis

by HPLC and/or LC-MS, however, as no tools exist for genetic engineering in *A. aculeatus* it is difficult to link genes to metabolites and explore silenced biosynthetic pathways.

Although the annotated genome of *A. aculeatus* has been available for more than two years¹, there are no reports on genome mining for identification of genes that are involved in secondary metabolism. In this study I have used *A. aculeatus* as a model for AMA1-based (non-integrative) gene expression, and shown that AMA1-based plasmids are maintained in *A. aculeatus*, that the *A. nidulans* *PgpdA* promoter controls gene expression in *A. aculeatus*, and that constitutive expression of a transcription factor activates a biosynthetic pathway whose products have not before been reported (Chapter 12).

¹<http://genome.jgi-psf.org/Aspac1/Aspac1.info.html>

1.3 Genetic engineering

Genetic engineering is a term that covers all genetic manipulations in a given organism e.g. gene deletion, gene disruption, gene insertion (random or site-specific), promoter exchange, silencing by RNAi, GFP-tagging etc. As the field of genetic engineering is so broad, it is not possible to apply one single method that will work for all possible disciplines. For instance, what might be a powerful method for gene deletion, may not be applicable for GFP-tagging. Therefore it is important to have a large available toolbox with a broad range of tools to choose from (Figure 1.5).

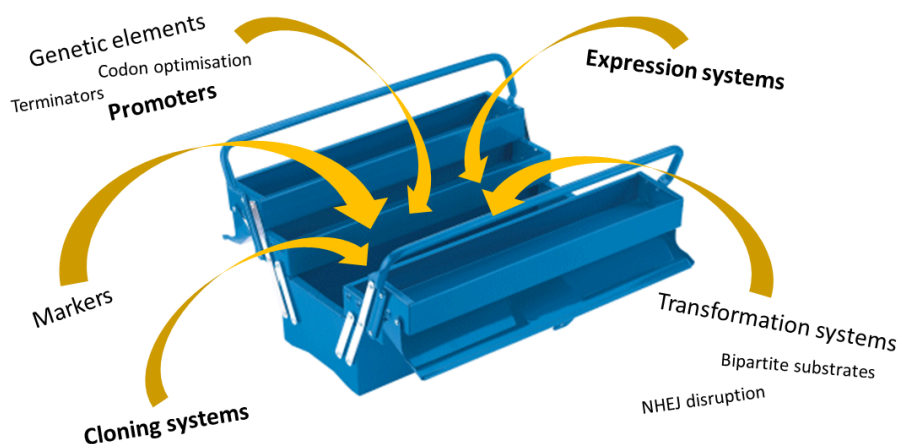


Figure 1.5: The toolbox for genetic engineering. Elements in bold have been primary foci of this study.

Such a toolbox should contain very basic elements such as promoters, terminators, and selection markers. Furthermore, it should contain a system for generation of transformation substrates, whether it be PCR-based or clonal. Last, but maybe most important, the toolbox should contain a transformation system for the organism in which the genetic engineering will be carried out. In this section I will briefly introduce the tools that are currently present in the *A. nidulans* and *A. niger* toolboxes. Subsequently, I will introduce the main vector construction approach used throughout this study: uracil excision (UE) cloning, and the present gene expression platform for *A. nidulans*.

1.3.1 Tools for genetic engineering in *Aspergillus nidulans*

During the past 30 years many attempts have been carried out in order to establish a platform for genetic engineering in Aspergilli, primarily *A. nidulans*, which was genetically well-established in the 1980s due to the preceding work of Pontecorvo. It was during the 1980s that most of the currently used genetic markers were described and characterized: *argB*, *pyrG*, *argB*, *trpC*, *riboB*, *amdS*, and *hph* [Buxton et al., 1985, Ballance et al., 1983, Yelton et al., 1984, Oakley et al., 1987, Tilburn et al., 1983, Cullen et al., 1987]. This decade also provided the two most widely used promoters in *A. nidulans*: the alcohol-inducible *PalcA* and the constitutive *PgpdA* [Lockington et al., 1985, Punt et al., 1990, 1991]. Recently, the repertoire of promoters was expanded

by the sucrose-inducible *PsucA* [Zuccaro et al., 2008, Roth and Dersch, 2010] and the doxycycline-sensitive tunable inducible O7 promoter, known as the TetON/OFF system [Meyer et al., 2011]. In this study I have primarily used the *PgpdA* promoter, as a strong and constitutively active promoter is an advantage both for testing new tools, and for expression of non-lethal biosynthetic genes. Additionally, I have used the O7 promoter for tunable over-expression of *A. nidulans* phosphatases and transcription factors (Chapter 6).

One of the very important tools that exists for *A. nidulans* is the use of color markers. The wild-type *A. nidulans* produces green conidia, however, the pathway for producing the green melanin pigment consists of two individual steps: synthesis of the yellow pigment YWA1 by the WA PKS (encoded by *wA*) [Watanabe et al., 1999] and polymerization of YWA1 by the laccase YA (encoded by *yA*) into green melanin (see Figure 1.6). Deletion of the *yA* gene results in a yellow strain and deletion of *wA*, which is epistatic to *yA*, results in a white strain. Hence, use of these color markers can help identify diploid strains, heterokaryons, gene targeting efficiencies and what not. In this study I have used the color markers for verifying expression from the newly developed AMA1 expression platform (Section 3.1.4), and for development of a background-free transformation system for gene expression (Section 3.1.5).

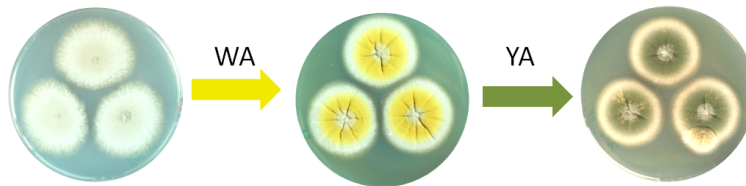


Figure 1.6: The pathway towards green melanin in *A. nidulans*. The WA PKS synthesizes yellow pigment YWA1, which is polymerized into green melanin by the YA laccase.

The annotated genome sequence of *A. nidulans* was released in 2005 [Galagan et al., 2005], and the availability of this sequence allowed for semi-large scale gene deletions e.g. the genome-wide phosphatase deletion study (27 genes) [Son and Osmani, 2009], the PKS deletion study (32 strains) [Nielsen et al., 2011], and the recent kinase deletion study (127 strains) [De Souza et al., 2013].

A very important tool, that made these semi-HTP strain constructions possible was the deletion of *nkuA*, a KU70 homologue, resulting in >90 % gene targeting efficiencies [Nayak et al., 2006, Nielsen et al., 2008]. NkuA is responsible for non-homologous end-joining (NHEJ), which is the most common DNA double strand break repair mechanism in *A. nidulans* [Nielsen et al., 2006]. This mechanism is also employed for random integration of transformation substrates, why screening for targeted integration events can be quite cumbersome in an *nkuA*⁺ strain.

In this study I have focused on assessing and improving tools for *A. nidulans* that are important for HTP genetic engineering: efficient cloning vectors (section 3.1.5) and efficient transformation systems (section 3.1.2). For future work, it is important not just to apply efficient methods for cloning and transformation, but also to have a wide available selection of genetic elements: promoters, terminators, markers, integration

sites etc. I have tried to expand the pool of available constitutively expressed *A. nidulans* promoters by analyzing expression data and testing selected promoters using a qualitative β -galactosidase assay, resulting in five novel strong promoters, of which one is a short bidirectional promoter (section 3.1.1). Lastly, I have expanded the repertoire of integration sites with five new sites that provide varying expression levels, making a total of eight available defined integration sites for *A. nidulans*, including the *yA* locus [Chiang et al., 2013] (section 3.1.2).

1.3.2 Tools for genetic engineering in *Aspergillus niger*

In the beginning of my PhD studies, I worked with a genetically wild-type fermentation strain of *A. niger*, ATCC1015, as this was the strain used in the fermentation activities at our research center and it was genome sequenced by the DOE Joint Genome Institute [Andersen et al., 2011]. As my collaboration with the analytical chemists at the research center expanded, I included several other strains of *A. niger* which had been isolated in various ecological environments. During my studies I engaged in a collaboration with Pacific Northwest National Laboratory (PNNL, Richland, WA, USA), which also had ongoing genetic activities with *A. niger*. They had prepared a derivative of ATCC1015, *KB1001*, which was *pyrG*⁻ and *kusA::pyrG* [Chiang et al., 2011]. The *kusA* gene is the homologue of *nkuA* in *A. nidulans*, why deletion of *kusA* allows for highly efficient gene targeting in *A. niger* (>80 % compared to 7 % in the wild-type) [Meyer et al., 2007]. The *kusA::pyrG* cassette in *KB1001* is a transient construction, so *pyrG* can be looped out and restore *kusA* to wild-type after genetic alterations have been made, as reported for *A. nidulans* [Nielsen et al., 2008]. I have used this strain for preparation of *A. niger* mutants, unless otherwise stated.

A disadvantage of using this strain is that the *pyrG* marker is used for making the *kusA* disruption and thus it cannot be used as a selection (and counter-selection) marker for iterative transformation steps. However, the ability to restore *KusA* after transformation is highly valued, as this protein is also involved in telomere maintenance [Fisher and Zakian, 2005, Palmer et al., 2010], which is crucial to the integrity of the strain. As the strain does not have any other auxotrophies, it is necessary to use dominant markers conferring resistance to e.g. hygromycin B or phleomycin which is a toxic and mutagenic substance. Furthermore, it is not possible to loop out the markers after transformation, why these markers can be used only once. The hygromycin phosphatase (*hph*) from *E. coli* confers resistance to hygromycin B, to which *A. niger* is naturally sensitive [Punt et al., 1987], and this setup has been successfully implemented for *A. niger* *KB1001* at PNNL. Therefore, I have used this marker for gene targeting in *A. niger* throughout my studies.

1.4 Uracil excision cloning

Over time, many different approaches for substrate generation have emerged, enabling construction of virtually any nucleic acid sequence one may wish. Basically, three options are available: PCR-based substrate generation, traditional cloning, or a combination of the two. Each method has advantages and disadvantages,

and thus the choice depends on what type of construct one wishes to make, and how many. In table 1.1 I have listed the methods that I believe are presently most commonly used, stating what I believe are their advantages and disadvantages. Generally, the use of PCR will always be mutagenic to some extent, but it allows for a flexibility in vector design that is absolutely essential in our time.

Based on the assessment of currently available methods for substrate generation, I have chosen to use the uracil excision (UE) cloning methods for generation of substrates for genetic engineering throughout this study. I have listed two types of UE cloning: PCR-based [Geu-Flores et al., 2007] and combined cassette/PCR [Nour-Eldin et al., 2006]. The PCR-based approach is solely based on fusion of PCR fragments for vector construction i.e. *E. coli* origin of replication and *E. coli* selection markers are also present as PCR fragments. The combined cassette/PCR approach uses a cassette, which contains recognition sites for a restriction endonuclease and a nicking enzyme. When digested, the cassette is opened and long single-stranded overhangs allow for fusion with PCR fragments. Thus several PCR fragments can be fused and inserted into the cassette in a single reaction. This will be described in detail below.

Method	Advantages	Disadvantages
(Fusion)PCR	Flexible	Inefficient
	Fast (<1 day)	Not robust
	Seamless fusion	Mutagenic
	Independent of restriction enzymes	
Traditional cloning	Not mutagenic	Inefficient
		Inflexible
		Leaves scars
		Screening laborious
Combined PCR/cloning	Flexible	Inefficient
		Leaves scars
		Mutagenic
		Screening laborious
In-Fusion®	Flexible	Low efficiency for multiple
	Seamless fusion of multiple fragments	fragments
	Independent of restriction enzymes	Mutagenic
Gibson Assembly®	Flexible	Low efficiency for multiple
	Seamless fusion of multiple fragments	fragments
	Independent of restriction enzymes	Mutagenic
		Expensive reagents
Gateway TM	Robust	Inflexible
	Efficient	Leaves large scars
	Suitable for HTP	Expensive reagents
Uracil excision, PCR	Very flexible	
	Robust	
	Efficient (for <5 fragments)	Mutagenic
	Seamless fusion of >10 fragments	Expensive primers
	Independent of restriction enzymes	
Uracil excision, combined cassette/PCR	Flexible	
	Robust	
	Efficient (for <5 fragments)	Mutagenic
	Seamless fusion of >10 fragments	Expensive primers
	Suitable for HTP	

Table 1.1: Assessment of current methods for substrate generation. Gibson Assembly®Master Mix and USERTM Reagent are products of New England Biolabs (MA, USA), GatewayTM is a product of Life Technologies (former Invitrogen), and the In-Fusion®Cloning System is a product of Clontech Laboratories, Inc.

UE cloning is also known as USERTM cloning, developed by New England Biolabs (Ipswich, MA, USA). Basically, the principle of UE cloning is to integrate a uracil into the PCR fragment, which will then be excised to form a long single-stranded overhang on the PCR fragment (see Figure 1.7). This single-stranded overhang will anneal to a complementary overhang and be ligated by the native ligase of the *E. coli* host after transformation.

Thus three reagents are needed in order to perform UE cloning: primers with uracil bases incorporated, a polymerase which does not stall at uracil bases, and the USERTM reagent containing enzymes for excision of the uracil base. Uracil-containing primers are available from most nucleotide suppliers, although at a higher price ($\approx 20\text{--}100\%$) than ordinary PCR primers. However, the appertaining polymerase, *PfuX7*, is freely available as an *E. coli* strain expressing the His-tagged polymerase [Nørholm, 2010], and the USERTM reagent is available from New England Biolabs at a very reasonable price ($\approx \$ 1/\text{reaction}$, July 2013). In comparison, the price for Gibson Assembly® reagent is $\approx \$ 15/\text{reaction}$ (New England Biolabs, July 2013).

The USER enzyme is a mixture of Uracil DNA glycosylase (UDG) and DNA glycosylase-lyase Endonuclease VIII (DGLE). UDG catalyzes cleavage of the glycosidic bond between the (deoxy)ribose and the uracil of the uridine in a RNA (or DNA) molecule [Lindahl et al., 1977, Lindahl, 1982]. This reaction excises the uracil base, but leaves the phosphodiester backbone intact (Figure 1.7). Subsequently, the DGLE breaks the phosphodiester backbone at the 3' and 5' ends of the abasic site so the deoxyribose is released [Melamede et al., 1994, Jiang et al., 1997]. The strength of the hydrogen bonds between the bases in the tail is not sufficient for keeping the two strands together, and thus the remaining bases of the tail will release from the double strand, generating a single-stranded overhang.

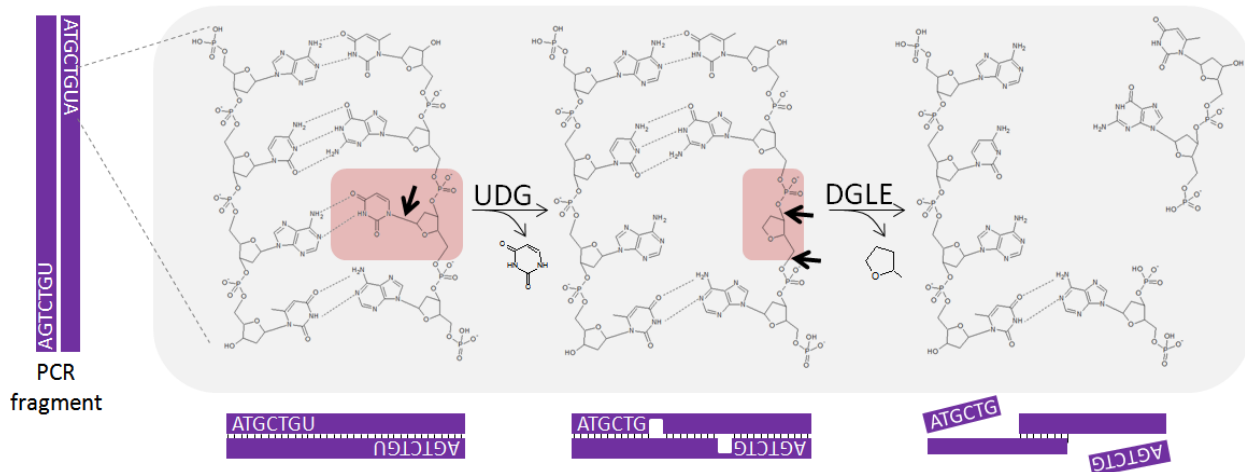


Figure 1.7: Generation of single-stranded overhangs by uracil base excision. The UDG cleaves the glycosidic bond between the deoxyribose and the uracil, releasing the uracil without disrupting the backbone. Subsequently the DGLE cleaves the backbone while releasing the deoxyribose. Hydrogen bonds in the short tail are too weak to keep the strands together, and the strands will disassemble, creating a 6 nt single-stranded overhang.

1.4.1 IS1 expression platform

UE cloning method was previously used to develop a series of cloning vectors for gene expression in *A. nidulans* [Hansen et al., 2011]. These vectors are all designed for integration into a defined and characterized site at the genome, designated Integration Site 1 (IS1) (Figure 1.8). The vectors contain one of two different promoters (*PgpdA* or *PalcA*) and one of two markers (*argB* or *pyrG*). All vectors contain 2 kb targeting sequences, TS1 (upstream) and TS2 (downstream), for targeted integration into the correct site.

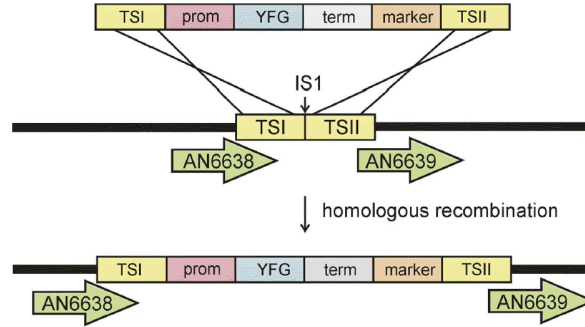


Figure 1.8: Integration into IS1, from [Hansen et al., 2011]. Linearized plasmid for gene expression contains an expression cassette comprising promoter, gene, terminator, and selection marker. The cassette is flanked by targeting sequences for targeted integration into IS1 by homologous recombination.

All vectors for this platform contain a UE-cassette between the promoter and terminator, in which any gene can be inserted that have compatible uracil-containing tails. Figure 1.9 describes the procedure for inserting a gene into the UE-cassette: the vector is prepared by digestion with a restriction endonuclease (AsiSI), and a nicking enzyme (Nb.BtsI) which will create long single-stranded overhangs. The gene is amplified by PCR, using primers that contain a predefined uracil-containing tail, complementary to the overhangs of the digested vector. The digested vector and the PCR product are mixed and incubated with the USERTM enzyme, which will remove uracils and expose single-stranded overhangs that can anneal to the single-stranded overhangs of the vector. Subsequently, the mixture is transformed into chemically competent *E. coli*, where the two fragments are ligated by the native *E. coli* T4 ligase. Correct plasmids are linearized and transformed directly into *A. nidulans*.

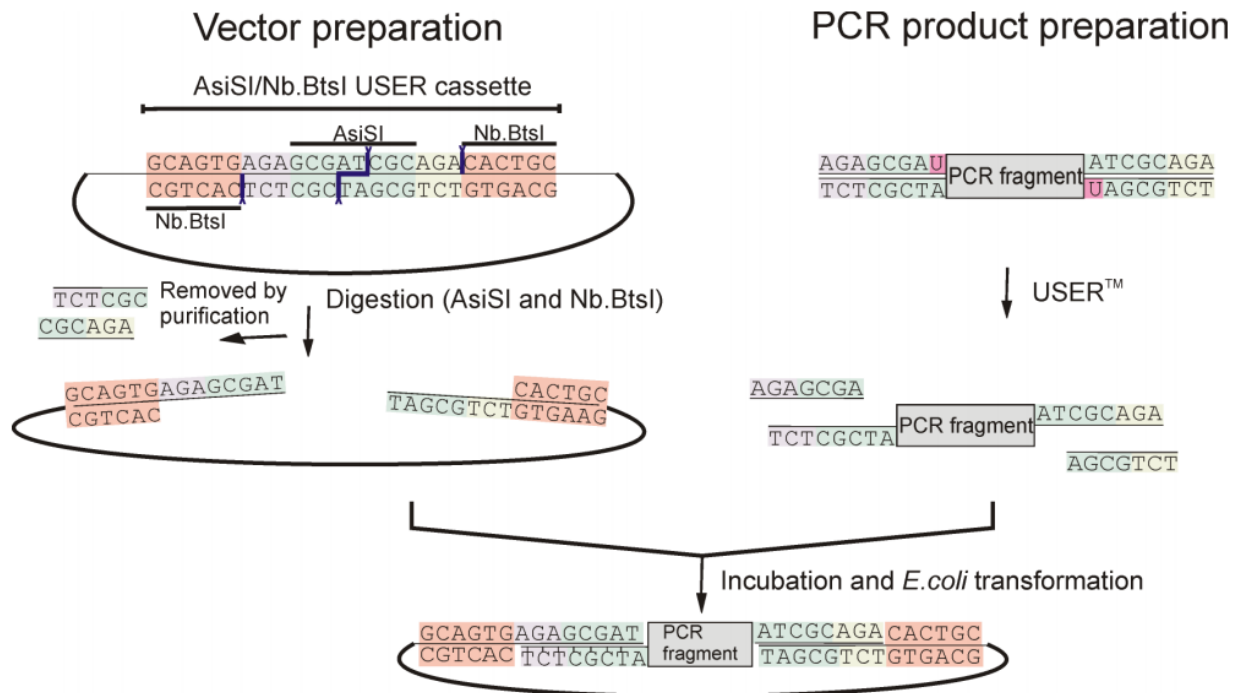


Figure 1.9: Principle of UE cloning, from [Hansen et al., 2011]. The vector is prepared by digestion with *AsiSI* and *Nb.BtsI* to generate single-stranded overhangs. PCR product is amplified, mixed with prepared vector, and incubated with *USER™* enzyme to generate complementary overhangs, which anneal to the overhangs of the vector.

1.5 Secondary metabolism in *Aspergillus*

The term *secondary metabolism* can be perceived in many ways, and many compounds can be regarded as both primary and secondary metabolites. In this thesis, I will use the term *secondary metabolite* only for metabolites that do not exist in primary metabolism, and *primary metabolism* comprising all reactions and metabolites that are essential for maintaining survival and growth in a *non-competitive* environment.

With respect to the above definition, secondary metabolites are organic molecules that are produced, in addition to the primary metabolites, by a given organism. Secondary metabolites are not directly required for sustaining life, but they may have essential roles within the cell e.g. protection against UV-light (pigments), warfare against competitors for nutrients (antibiotics, antifungals, insecticides) or intra-species signaling. Many secondary metabolites have been isolated and structure elucidated, however, for most of them their biological functions are yet unknown.

Primarily, three classes of secondary metabolites exist in filamentous fungi: polyketides (PKs), nitrogen-containing non-ribosomal peptides (NRPs)/alkaloids, and terpenoids (examples are shown in Figure 1.10). PKs are synthesized in long chains of adjoining ketide units by a polyketide synthase (PKS), which is similar to the fatty acid synthase and employs the same mode of action. NRPs are, as the name implies, peptides that are not synthesized by a ribosome, but by a non-ribosomal peptide synthase (NRPS) which selects and joins specific amino acids via peptide bond formation. Terpenoids are hydrocarbon molecules which are composed by C₅ isoprene units. In the following I will describe synthesis of PKs by the PKS, the structure of the PKS, and organization of a typical PK gene cluster. As my study has dealt primarily with biosynthesis of PKs, I will not go into further detail with synthesis of NRPs, alkaloids, or terpenoids.

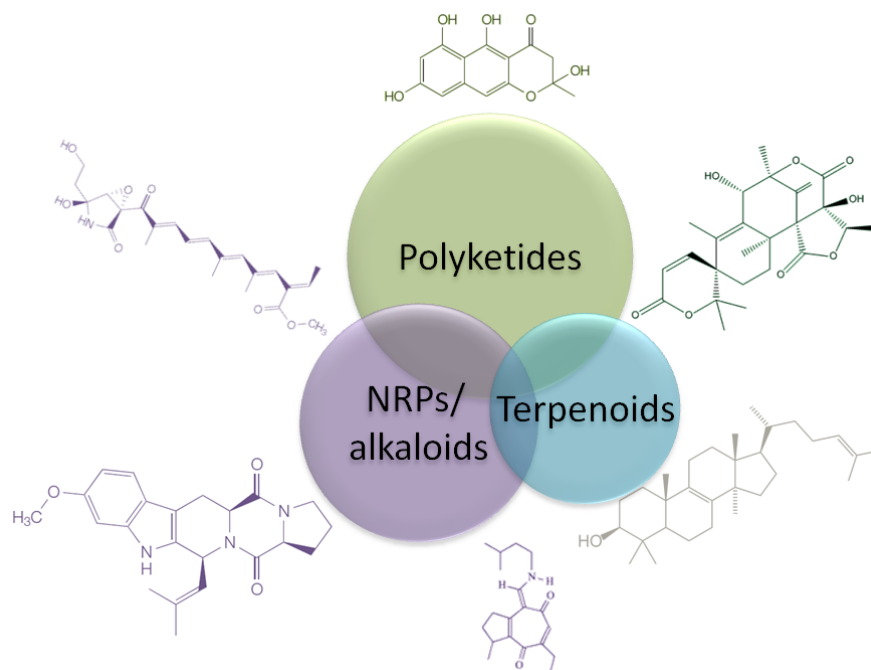


Figure 1.10: Fungal secondary metabolites normally comprise three different classes: polyketides (PKs), non-ribosomal peptides (NRPs), and terpenoids. A compound may be hybrids of two (or all) different classes.

1.5.1 Fungal polyketides

Polyketides are a group of secondary metabolites from bacteria, fungi, plants, and animals, which are predicted to have a great potential as antibiotics and natural colourants. In fungi, PKs are the most abundant group of secondary metabolites [Keller et al., 2005]. Some of the best known fungal polyketides, produced by type I iterative PKS, are the cholesterol lowering commercial drug *lovastatin* and the highly carcinogenic *aflatoxin*, which is produced by e.g. *Aspergillus flavus*. Although PKs are very diverse in structure and function, they are synthesized in the same manner, from the same pool of extender units, using the same types of enzymes.

These enzymes can either exist as a system of individual domains (type II polyketide synthase) or the domains can be assembled into a multifunctional multi-domain enzyme structure (type I polyketide synthase) (Figure 1.11). The type I PKS can further be divided into two categories: modular or iterative [Katz and Donadio, 1993]. Modular PKSs comprise a number of *modules*, where one module contains the set of enzyme domains that are needed for each round of elongation (addition of one extender unit) [Donadio et al., 1991]. As modular PKSs contain a domain for each single activity, and even in an orderly manner, it is possible to predict the PK product of the modular PKS.

Iterative PKSs, however, contain only a single module, which is re-used in an iterative fashion until the product is finished. In contrast to modular PKSs, the products of iterative PKSs cannot easily be predicted [Cox and Simpson, 2009]. In fungi, the PKs are most often synthesized by a type I iterative PKS, why it is

very difficult to predict which compound(s) that are products of the PKSs. Much research have been done in an effort to unveil the mechanisms of determining chain length, and when to turn on specific domains, but very little is still known about these mechanisms.

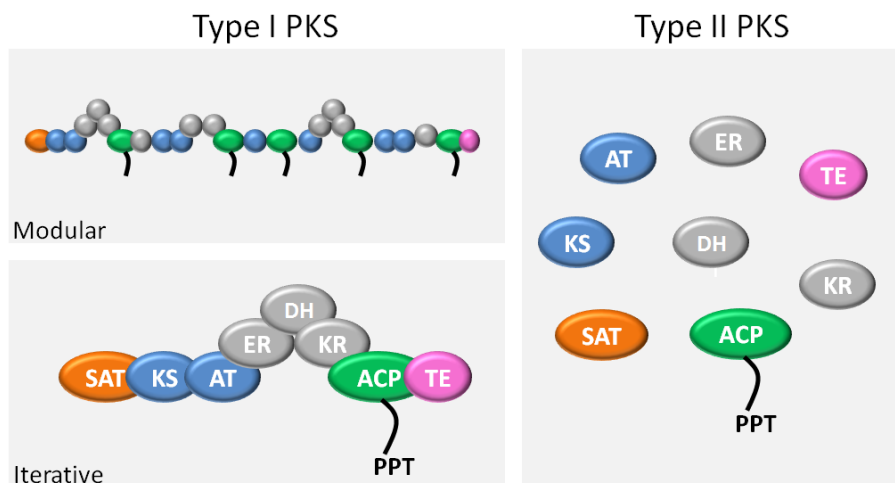


Figure 1.11: Types of polyketide synthases (PKSs). The type II PKS exists as discrete enzymes which form a complex to synthesize the PK. The type I PKS is one single multidomain enzyme, and can be either *modular* or *iterative*. The modular PKS contains several modules, each module containing the necessary domains for one round of elongation. Iterative PKSs comprise only one module, which is re-used in an iterative fashion until the final PK has been synthesized.

Domains of the PKS

Basically, three domains are essential for synthesis of a PK by the PKS: the β -ketoacyl synthase (KS) domain, the acyltransferase (AT) domain, and the acyl carrier protein (ACP) domain. These three domains alone can recruit building blocks and synthesize the PK chain, but most often the PKS contains additional domains. All known domains from type I iterative PKSs are summarized in Table 1.2.

PKSs are divided into three groups based on their potential to reduce polyketide backbones: non-reducing PKSs (NR-PKSs), partly reducing PKSs, and highly reducing PKSs. The NR-PKSs do not contain any of the reducing domains (KR, DH, or ER - see Table 1.2), whereas the reducing PKSs contain either KR or KR/DH (partly reducing), or all three domains (highly reducing). The NR-PKSs are known to produce cyclic PKs, which is not the case for reducing PKSs, and thus the NR-PKSs contain two additional domains which have not been reported for any reducing PKS: the *product template* (PT) domain and the *cyclization* (CYC) domain. The PT is responsible for correct folding of the PK, and the CYC domain is often a combined cyclization/release mechanism [Crawford et al., 2008, 2009].

Methyl transferase (MT) domains are found in all types of PKS, and they catalyze the addition of a methyl group to the growing PK chain, using S-adenosylmethionine (SAM) as a methyl donor [Cox and Simpson, 2009].

One activity that is very important in the aspect of producing PKs, is *release* of the final PK from the PKS. This can be accomplished in many ways, and is poorly understood. It is known that the CYC and thioesterase (TE) domains can cleave the PK chain from the PKS by hydrolysis [Weissman, 2008, Schroeckh et al., 2009]. It was also recently shown that the dehydratase (DH) domain of the *A. terreus* 6-methylsalicylic acid (6-MSA) synthase was in fact a *thioester hydrolase* (TH) domain that mediated release of 6-MSA from the PKS [Moriguchi et al., 2010]. Another known release mechanism is the *thioester reductase* (TR), but this mechanism is more pronounced for release of NRPs from the NRPSs. Additionally, many PKSs rely on discrete enzymes in the gene cluster for mediation of release e.g. the fumonisins PKS relies on Fum8 for product release [Gerber et al., 2009].

Minimal set of domains	
KS	(β -ketoacyl synthase) Catalyzes synthesis of the polyketide chain by joining building blocks in a Claisen-like condensation reaction. KS is similar to the condensing enzyme (CE) in the fatty acid synthase complex [Weissman, 2008].
M/AT	(Malonyl/acyl transferase) Catalyzes the loading of the starter unit, as well as transfer of acyl from acyl-CoA to the growing ketide chain [Cox, 2007].
ACP	(Acyl carrier protein) Similar to ACP in fatty acid synthesis. ACP is a cofactor that facilitates elongation of the polyketide chain by carrying malonate units and transiently holding the growing ketide chain [Cox, 2007, Weissman, 2008]. This domain is also known as the phosphopantetheine (PPT) attachment site as it contains a PPT prosthetic group which is required for covalent attachment to the acyl starter unit [Long et al., 2002, Kroken et al., 2003, Cox, 2007, Weissman, 2008].
Reductive domains	
KR	(Keto reductase) Catalyzes the reduction of a ketone (C=O) into a hydroxyl group (C-OH). This domain is only present in PKSs that are reducing or partially reducing.
DH	(Dehydratase) Catalyzes dehydration of hydroxyl groups (obtained by the KR), into an enoyl group (-C=C-).
ER	(Enoyl reductase) Catalyzes the reduction of enoyl groups (obtained by the DH) into a saturated carbon atom (-C-C-). This domain is only present in highly reducing PKS.
Other domains	
PT	(Product template) Mediates the cyclization of aromatic polyketides from non-reducing PKSs. This domains is believed to only be present in non-reducing PKSs [Crawford et al., 2008, 2009].
CYC	(Cyclization) Catalyzes cyclization of a reduced polyketide. This domain is often a combined CYC/TE domain for simultaneous cyclization and release.
TE	(Thioesterase) Mediates the release of the complete polyketide from the enzyme. This domain is believed to be very substrate specific [Weissman, 2008, Schroeckh et al., 2009].
TH	(Thioester hydrolase) Newly assigned domain in 6-MSAS (<i>Aspergillus terreus</i>) catalyzing product release by hydrolysis [Moriguchi et al., 2010].
MT	(Methyl transferase) Catalyzes the addition of a methyl group to the (poly-)ketide, using S-adenosylmethionine (SAM) as a methyl donor [Cox and Simpson, 2009].

Table 1.2: Overview of domains in the iterative type I PKS.

1.5.2 Synthesis by the type I iterative polyketide synthases

As briefly mentioned above, the type I iterative polyketide synthases differs from type I modular polyketide synthases, as they are multidomain enzymes that consists of only one module, whereas type I modular polyketide synthases consist of several modules. Thus, this one module is re-used a certain number of times (3-20 iterations), depending on the length of the polyketide that is to be synthesized. The PK chain is normally synthesized from acyl-CoA and elongated with malonyl-CoA units [Schumann and Hertweck, 2006] (Figure 1.12), however the PKS may employ alternative starter units.

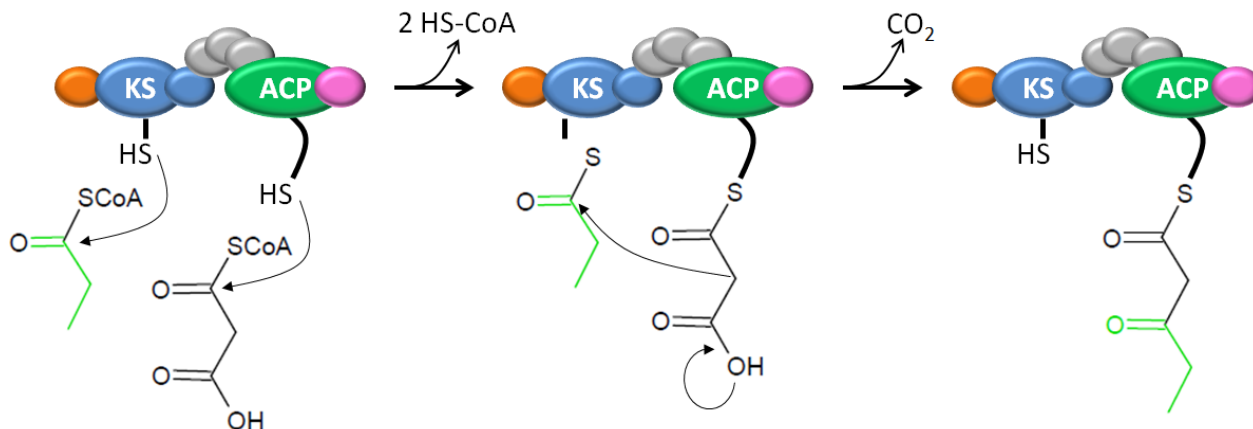


Figure 1.12: Schematic illustration of synthesis by a fungal iterative PKS. Acetyl-CoA is loaded onto the KS domain, and the malonyl-CoA extender unit is loaded onto the ACP. Malonyl-ACP and acetyl-KS condense in a Claisen-like reaction, extending the acetyl starter unit with one ketide and releasing CO_2 .

Synthesis of a PK always starts with *activation* of the PKS. Activation is accomplished by attachment of a *4'-phosphopanthetheinyl* (PPT) prosthetic group to the ACP domain by a dedicated enzyme, the PPTase. The PPT acts as the "arm" that carries the growing PK and exposes it to the domains of the PKS. A *starter unit acyltransferase* (SAT), if present in the PKS, will choose the starter unit, otherwise an acetyl-CoA will serve as the starter unit. The starter unit (-CoA) will be loaded onto the KS (see Figure 1.12), while the AT recruits the extender unit (most often malonyl-CoA) and attaches it to the ACP. The acetyl-CoA starter unit and the malonyl-CoA will condense in a decarboxylative Claisen-like reaction by a nucleophilic attack on the carbonyl of acetyl-CoA, resulting in a β -keto thioester [Schumann and Hertweck, 2006, Sta]. Elongation of the PK chain by addition of ketide units will continue until the final length has been reached. At present, it is not known how the PKS controls the length of the PK.

During synthesis, the PK chain may be modified by addition of methyl groups or by reduction of the ketides (Figure 1.13). There are three degrees of reduction for the ketide: reduction of the carbonyl into a hydroxyl (by the KR domain), dehydration of the hydroxyl into an enoyl (by the DH domain), and reduction of the enoyl double bond into a single bond (by the ER domain). Even though a PKS contains all three

reductive domains they may not be used in every iteration, why the final PK from a highly reducing PKS may contain a mixture of intact ketide units, hydroxyls, double bonds and saturated carbons. Thus the presence of these domains add to the complexity and diversity of different polyketides.

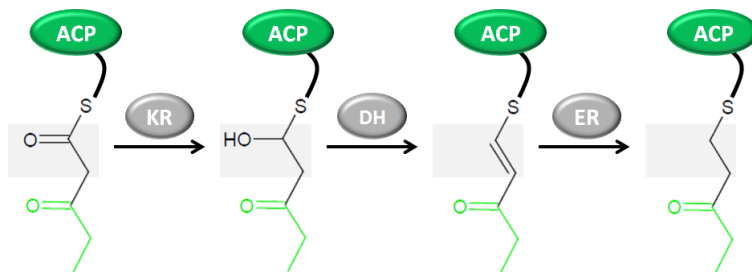


Figure 1.13: Reduction of the ketide unit. Ketoreductase (KR) reduces the keto-group to a hydroxyl-group, dehydratase (DH) dehydrates the hydroxyl-group to a double bond, and the enoyl reductase (ER) reduces the double bond to a single bond.

1.5.3 Secondary metabolite gene clusters

As mentioned above, the presence of additional methylation and reduction domains in a PKS increases the complexity of the PK and the pool of different PK varieties. In addition to these peri-synthetic modifications, PK are often also modified post-synthetically i.e. after release from the PKS. Post-synthetic modification may also comprise methylation and/or reduction, but often the changes are more drastic e.g. reorganization of the entire backbone as is the case for austinols [Lo et al., 2012]. The genes encoding the enzymes that are required for post-synthetic modification of PKs are most often located adjacent to the PKS gene, and there are rarely any unrelated genes in between. As the genes are clustered together in the genome, they are referred to as a *gene cluster*.

A PK gene cluster normally contains 5-15 genes, depending on the structure and nature of the final PK. Besides enzymes for modification of the PK, the gene cluster may also encode other proteins that are necessary for successful production of the PK. An important activity encoded by the gene cluster, is resistance to the PK. Many of the secondary metabolites are used to kill other organism in the environment and thus the compounds may also be toxic to the fungus itself. This problem can be solved by producing a protein that directly confers resistance to the toxic compound, or by producing pumps that will export the compound out of the cell. These activities will normally also be encoded by the gene cluster. A schematic of a typical secondary metabolite gene cluster is shown in Figure 1.14.

1.5.4 Transcription factors

As secondary metabolites are very expensive molecules to produce for the cell, and not essential for growth and survival in a non-competitive environment, the cell may employ strict regulation of a given PK in order to ensure that the compound is only produced when necessary. This regulation may be global or local. A global

regulation mechanism e.g. DNA acetylation, would affect production of many or all secondary metabolites in the cell, whereas a local regulation mechanism would only affect production of a single or maybe a few secondary metabolites within the cell. A local regulation mechanism can be obtained by employing specific *transcription factors* (TFs) that either facilitate transcription (*activators*) or obstruct transcription (*repressors*), and the genes that encode specific TFs are often located in the gene cluster (see Figure 1.14).

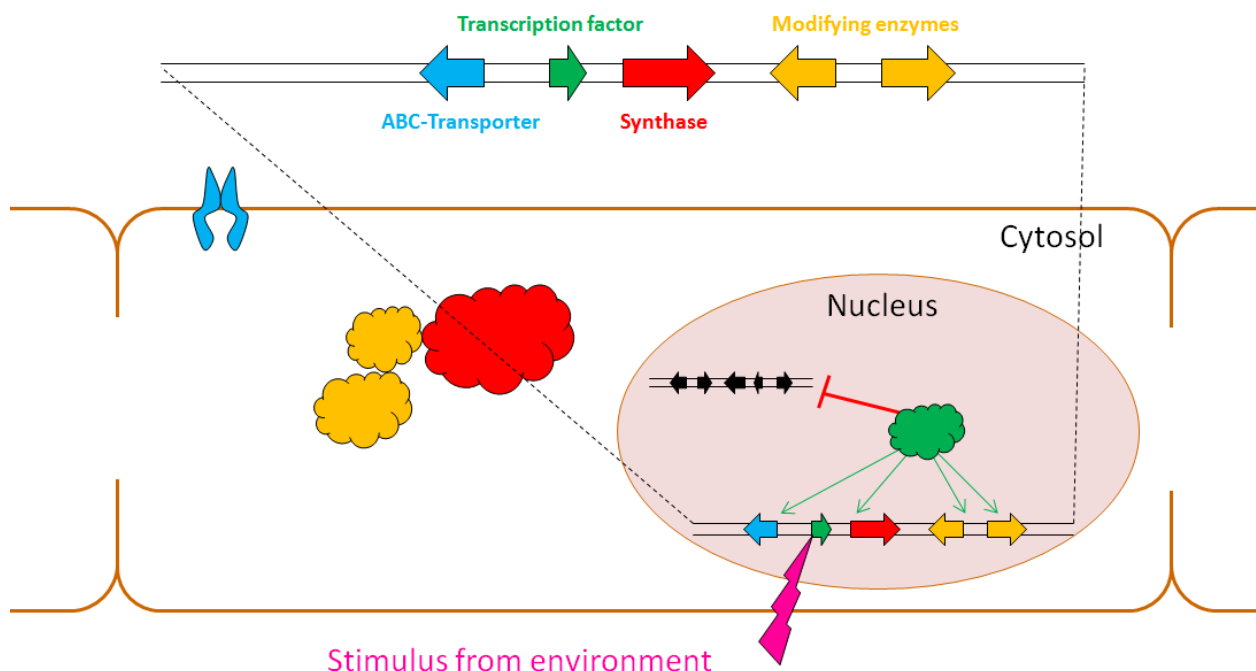


Figure 1.14: Schematic of a typical secondary metabolite gene cluster that is regulated by a TF. A stimulus from the environment activates the TF or triggers expression of the TF. The TF activates expression of all genes in the cluster, and simultaneously represses a second and different gene cluster (shown in black). The activated gene cluster produces a PKS and two tailoring enzymes that synthesize the final PK, and an ATP-binding cassette (ABC)-transporter for active transport of the PK out of the cell.

The first discovery of genetic regulation was in the study of the *lac* operon in *E. coli* [Jacob and Monod, 1961], but it was not until twenty years later that the *cis-acting regulatory elements* were discovered [Davidson et al., 1983]. A *cis-acting regulatory element* is a stretch of DNA that regulates expression of a gene, and it can be located upstream or downstream of the regulated gene, but it may also be located in an intron or even in an exon. The *cis-acting regulatory element* is the sequence to which the transcription factor (TF) binds, whether the TF acts as an activator or a repressor. A *trans-acting regulatory element* also refers to a stretch of DNA that regulates expression of a gene, however, in a different way. Normally, the *trans-acting regulatory element* comprises a DNA sequence that encodes a protein (e.g. a TF) or microRNA, which will affect regulation of one or several target genes.

The general role of TFs is to regulate the processes that take place within a given organism: nutrition uptake and degradation, production of secondary metabolites, growth, stress responses, and cell differentiation, just to mention a few. The number of TFs in a given organism is linked to the complexity of the organism, which is illustrated by the fact that *E. coli* contains 48 TFs, *S. cerevisiae* contains 169 TFs, and the human genome contains ≈ 2600 TFs [Blattner et al., 1997, Goffeau et al., 1996].

Sequencing and subsequent annotation indicates that the *A. niger* genome contains almost a thousand (944) TFs [Pel et al., 2007], while the *A. nidulans* genome merely contains 372 TFs [Galagan et al., 2005, Wortman et al., 2009]. In this study I have used TFs for elucidation of biosynthetic clusters, and I have systematically over-expressed *A. nidulans* TFs in an effort to perturb metabolism and unveil new metabolites (Chapter 6).

1.5.5 Heterologous expression of secondary metabolite genes in *Aspergillus*

The approach of expressing genes heterologously (in another organism) have been done for many years, primarily using bacteria or yeast as host organisms. These organisms are good choices for production of proteins or enzymes as they are well-characterized and easy to cultivate and manipulate genetically. However, when studying secondary metabolism, there are good reasons why one should use a filamentous fungus as host instead of bacteria or yeast. The primary reason is that filamentous fungi are natural producers of a wide range of secondary metabolites, why they already possess the molecular machinery that facilitates production thereof. Secondly, filamentous fungi are able to splice introns, which is not the case for bacteria and simple yeasts e.g. *S. cerevisiae*.

Many bacteria are genetically well-characterized and easy to cultivate, however, there are many reasons not to choose the bacterial cell factory for heterologous production of secondary metabolites. First, there is often a codon-bias that necessitates codon-optimization for efficient expression. Furthermore, access to required building blocks, correct protein folding, any glycosylation or other posttranslational modifications e.g. attachment of the PPT to the ACP domain of PKSs and NRPSs, may not be possible.

For the established cell-factory *S. cerevisiae* there are only a few successful examples of heterologous expression of fungal PKSs [Kealey et al., 1998, Zhou et al., 2010, 2012, Ishiuchi et al., 2012], and only one report on reconstruction of a biosynthetic pathway [Rugbjerg et al., 2013].

In contrast, there are many successful reports on heterologous expression of PKSs in *A. nidulans*, reviewed in [Schumann and Hertweck, 2006]. Some of the important advances was cloning and expression of the 6-MSA synthase from *A. terreus* [Fujii et al., 1996], production of lovastatin by co-expression of several genes [Kennedy et al., 1999], development of a systematic gene expression system [Hansen et al., 2011], and reconstruction of the entire geodin cluster from *A. terreus* [Nielsen et al., 2013] (Chapter 11).

Gene prediction

Successful expression of genes is dependent on correct *gene prediction* i.e. localization of the open reading frame (ORF) and possible exons and introns within the ORF. The ORF, which comprises the entire stretch of DNA to be transcribed, may also contain untranslated regions (UTRs) that can be located in both the 3' end and the 5' end of the gene. Prediction of putative genes in a genome is performed automatically by applying one or several different gene prediction algorithms, followed by manual inspection by experts. Even so, the complexity of eukaryotic genes makes it difficult to accurately predict the correct ORF, especially for filamentous fungi for which only a few species have been genome sequenced, and little research have been done to uncover the aspects of gene organization.

There are several web-based options for gene predictions, which can support the predictions provided by the genome databases. For *A. nidulans* I have used the predictions provided by the Aspergillus Genome Database (AspGD) (<http://www.aspergillusgenome.org/>), and for *A. niger*, *A. aculeatus*, and *A. carbonarius* I have used the predictions provided by the DOE Joint Genome Institute (JGI) (<http://www.jgi.doe.gov/>). When in doubt of correct prediction, I have used the two web-based gene predictors FGeneSH (Softberry, Mount Kisco, NY, USA) and AUGUSTUS [Stanke et al., 2004, Stanke and Morgenstern, 2005] for support.

Prediction of introns within an ORF is very important for heterologous expression of the gene. Although intron loss and gain in Aspergilli have been studied [Nielsen et al., 2004, Zhang et al., 2010, Torriani et al., 2011, Croll and McDonald, 2012], no research has been conducted on the ability to identify and splice introns from other organisms. Thus it would be optimal to use intron-free versions of genes that are expressed heterologously. Basically, there are two ways to remove introns in a gene: to amplify the gene from cDNA or to PCR amplify and fuse exons (Figure 1.15).

Transcription of a eukaryotic gene results in a pre-mRNA molecule that contains the entire ORF sequence and in addition, it is capped in the 5' end and a poly-adenylation signal (poly-A tail) is attached to the 3' end (Figure 1.15). The introns are then spliced out of the pre-mRNA, resulting in the mature mRNA that now only contains the coding sequence (CDS) with the 5' cap and the 3' poly-A tail. At this point, total mRNA can be extracted from the cell, and cDNA is produced *in vitro* by reverse transcription. The intron-free gene is then PCR amplified from the cDNA template using gene-specific primers. Another way to produce an intron-free gene is by fusion of exons (Figure 1.15). In this case, the predicted exons are PCR amplified using primers with complementary tails. The exons can then be fused by any available method e.g. fusion PCR, UE cloning, In-Fusion® etc.

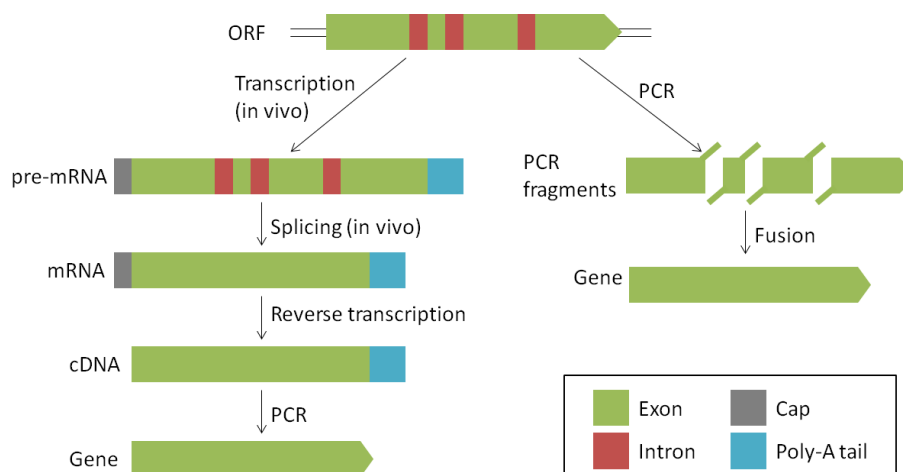


Figure 1.15: Approaches for generating intron-free DNA. Intron-free DNA can be generated by reverse transcription of mRNA into cDNA (left) or by fusion of exons (right). See text for details.

Each of the two approaches have advantages and disadvantages. The fusion approach is dependent on correct intron prediction in order to produce a functional protein, and the efficiency of the fusion method is dependent of the number of exons. If the gene contains only two exons the fusion may be possible by any method, but if the gene contains >10 exons, fusion may be difficult. The cDNA approach is the most accurate: the mRNA is spliced and assembled correctly, producing an intact and functional protein. The disadvantage of the approach is that it can be very difficult to obtain long cDNA fragments. In this study I have primarily expressed PKS genes, which are typically 7-10 kb in length. In my experience, it has been very difficult to obtain cDNA of this length, why I have not been able to use this approach in a robust manner. Combined with the risk of incorrect intron prediction, I have chosen to use the entire ORF for heterologous gene expression throughout this study. Thus I have assumed that introns and intron splicing has not evolved much within fungi of the same genus, and I hypothesize that *A. nidulans* is able to correctly splice introns from other *Aspergilli*.

1.6 Metabolite profile analysis

In this thesis *metabolite profile analysis* (MPA) refers to chemical analysis of wild-type or mutant strains by UHPLC-DAD-TOFMS, followed by data analysis (dereplication) to reveal which compounds are present in the sample. The MPA can be carried out manually or partly automated, and can comprise full analysis of all compounds (often more than 100 compounds/sample), or analysis of selected relevant compounds. In the following, I will describe just the basics of MPA, which will allow for an understanding of the obtained results.

The first step of MPA is to extract compounds from a given strain. As secondary metabolites are primarily expressed on solid media [Andersen et al., 2013], the strain that is to be analyzed is most often grown on agar

plates at an appropriate temperature. Compounds are extracted by cutting out a part of the agar/fungus and extracting it with solvent(s) e.g. ethyl acetate, methanol, or water. If the chemical properties of the compounds are right with respect to polarity etc., they will diffuse from the agar into the solvent. Thus, the choice of solvent depends on the compound that one wishes to extract. In this study I have primarily used a mixture of solvents, which will extract a wide range of secondary metabolites, and are therefore suitable for analysis of the entire metabolite profile [Smedsgaard, 1997].

After preparation of the extract it is processed on the UHPLC-DAD-TOFMS system, which comprises three different systems: the Ultra High Performance Liquid Chromatography (UHPLC) system that separates the compounds based on chemical properties (in this case: polarity), the Diode Array Detector (DAD) system that detects the compounds based on absorption of UV/Visible light, and the Time-of-Flight Mass spectrometry (TOFMS) system that separates the ions based on their m/z ratio, by measuring the time-of-flight (TOF) i.e. how long time it takes for the ion to reach the detector (Figure 1.16). Thus the output for each different compounds is the *retention time* (RT) on the column, the *UV/VIS spectrum*, and the *mass spectrum*, and in combination these features are almost unique for a given compound, or at least for a family of compounds.

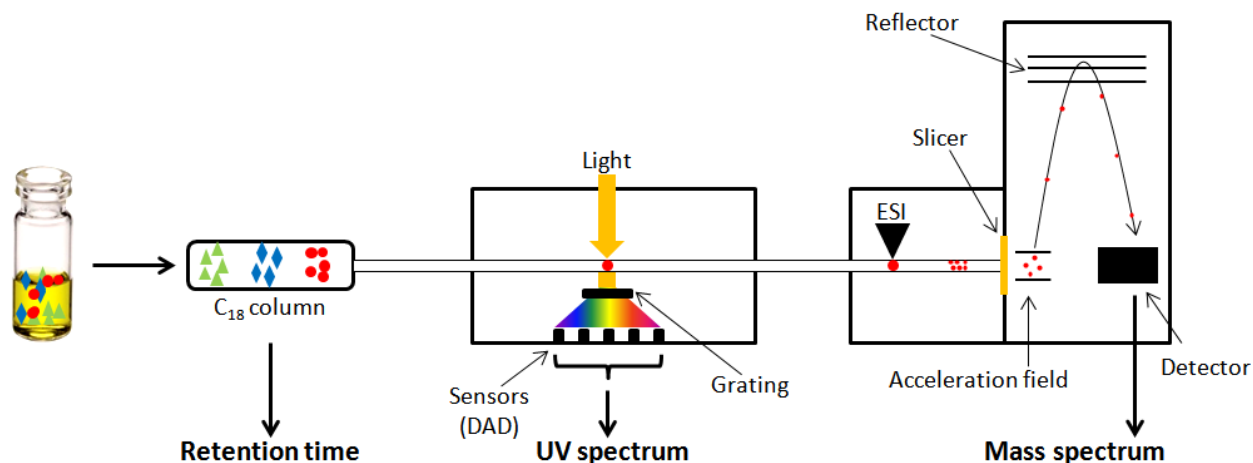


Figure 1.16: Principle of the UHPLC-DAD-TOFMS system. Extract is injected and compounds are separated by polarity on C_{18} column. Separated compounds (indicated by a red sphere) are subjected to a beam of light. Non-absorbed light is dispersed by the grating and detected by sensors (DAD). Subsequently, the compounds are dispersed into charged micro-droplets by the ESI and then enter the TOF in small "portions" via the slicer mechanism. The ions are accelerated in the acceleration field, and reflected back where the time-of-flight is measured by the detector.

1.6.1 Liquid chromatography

Liquid chromatography (LC), or *Ultra-High Performance Liquid Chromatography* (UHPLC) is a method for separating chemical compounds based on their chemical properties. In this study I have used a C_{18}

column (the stationary phase), which separates compounds based on polarity. The C_{18} molecules present in the column will interact with apolar compounds (e.g. hydrocarbon chains or steroids) but less so with polar compounds (e.g. sugars, amino acids). The *mobile phase* is a liquid solvent used for elution of the compounds. The mobile phase may be a mixture of different solvents, and the ratio of solvents may vary during the run.

Figure 1.17 shows an example of a *chromatogram*, which is the output from the UHPLC. The extract shown is from the *A. niger* reference strain KB1001, cultivated on YES agar plates for 5 days at 30°C. The runtime of the sample is 15 minutes, and each peak in the chromatogram represents detection of one or more specific compound(s). Thus, >100 compounds are detected in this sample. The intensity of a given peak may be a qualitative measure of the amount of compound in the sample, however, the intensity may also reflect the degree of ionization, which allows it to be detected.

As polar compounds do not interact strongly with the C_{18} column, these compounds will not be *retained* by the stationary phase and they will elute first. Apolar compounds will be retained longer and will have a higher *retention time* (RT). Figure 1.17 shows three compounds that are produced by *A. niger*: *pyranonigrin A*, which is a small very polar compound with a low RT (2.3 min); *aurasperone B*, which is a rather large and semi-polar compound with an intermediate RT (6.6 min); and *yanuthone D*, which is a large and apolar compound with a high RT (9.3 min).

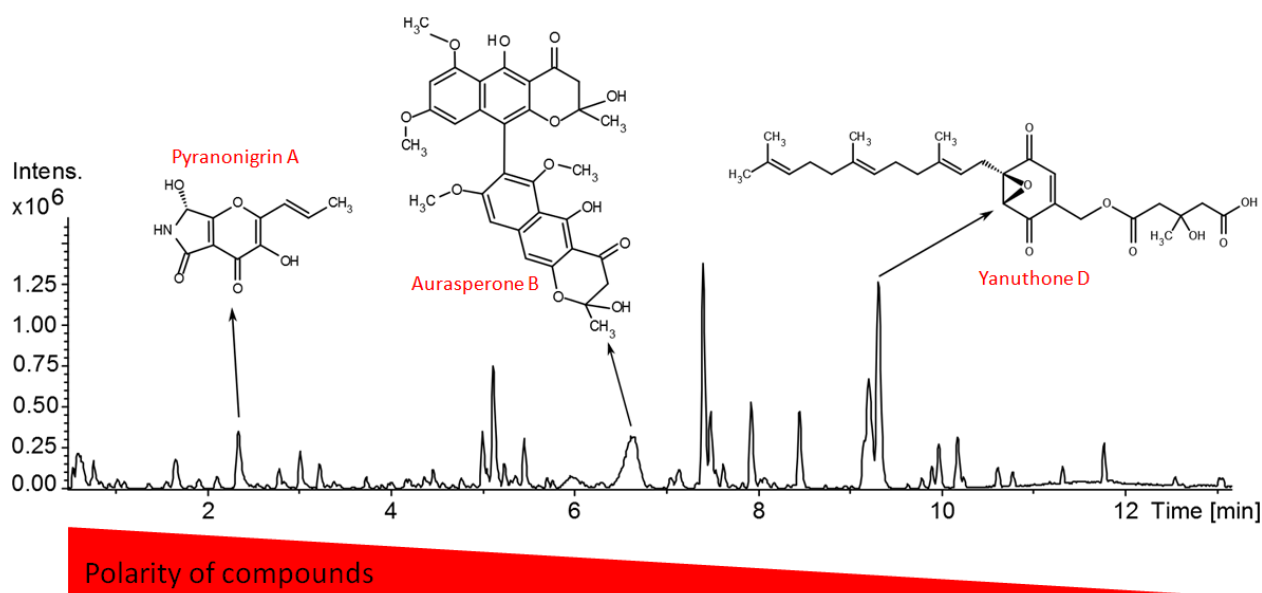


Figure 1.17: Base peak chromatogram (BPC) of *A. niger* extract. For details, see text.

1.6.2 Diode Array Detection (DAD)

The *Diode Array Detector* (DAD) is a method for detection of compounds based on their ability to absorb light. The DAD emits an array of UV/visible (UV/VIS) light (200-700 nm), and measures absorption of

light by the compound within the UV/VIS range. The output of this detection is the *UV/VIS spectrum*, with wavelengths on the x-axis and relative absorbance on the y-axis. Figure 1.18 shows the UV/VIS spectra of pyranonigrin A, aurasperone B, and yanuthone D. All three compounds have *chromophores* consisting of conjugated double bonds (marked in red), but the amount and position of conjugated double bonds determines which wavelengths that can be absorbed. The UV/VIS spectrum can thus be considered a finger-print of a compound family, as it is unique to the amount and organization of the chromophore.

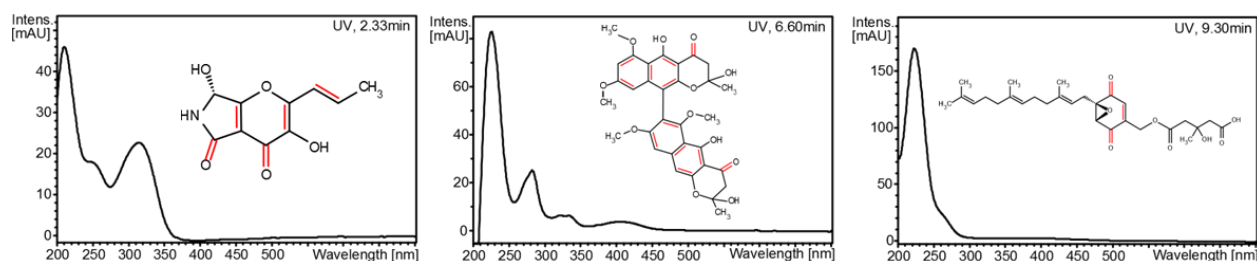


Figure 1.18: UV spectra of pyranonigrin A, aurasperone B, and yanuthone D. Conjugated double-bonds are shown in red.

1.6.3 Mass Spectrometry (MS)

Mass Spectrometry is a method for detecting compounds based on their masses. The basic principle of the analysis method is by formation of ions (positive or negative ionisation), and subsequent measuring the mass-to-charge ratio (m/z) of the resulting ions. The output of the MS is a *mass spectrum*, from which information on the monoisotopic mass of the compound can be deduced. The mass spectra of pyranonigrin A, aurasperone B, and yanuthone D are shown in Figure 1.19; m/z on the x-axis, and intensity on the y-axis.

The protonated compound is called the *pseudo-molecular ion*, and is designated $[M+H]^+$ (shown in green). Thus, the monoisotopic mass of the compound equals $m/z - H^+$, if there is only one charge on the compound. The remaining "sticks" in the mass spectrum are called *adducts* and are different ions of the compound. There are only a few possible adducts for each compound, why these are very useful in deducing which ion is the pseudo-molecular ion.

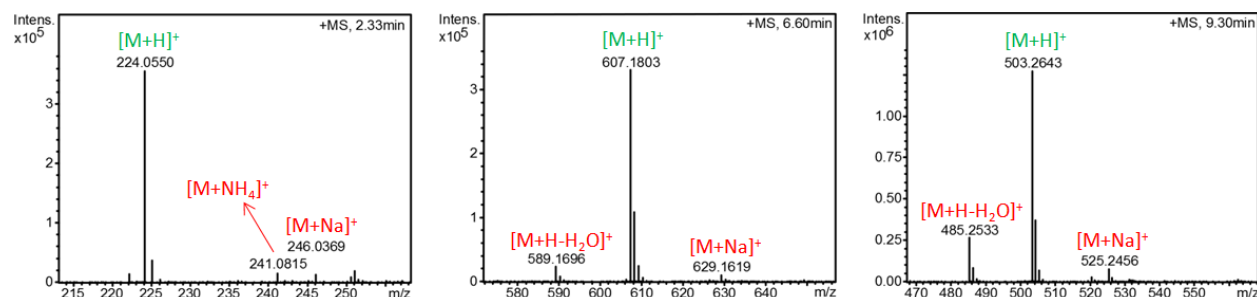


Figure 1.19: Mass spectra of pyranonigrin A, aurasperone B, and yanuthone D. The pseudo-molecular ion is shown in green, and adducts are shown in red.

The mass spectrum provides much more information than just the mass of a given compound. The time-of-flight MS (TOFMS) is so accurate (<1 ppm error) that the precise chemical composition may be deduced from the accurate mass alone. The mass spectrum thus also provides isotopic patterns, which reveal the number of carbon, chlorine, bromine, and sulphur atoms in the compound. Hence, the combination of UHPLC, DAD, and MS is very powerful for identification of compounds, both known and unknown.

Chapter 2

Overall results and discussion

2.1 Investigation of secondary metabolism in *A. nidulans*: a fungus with a large genetic toolbox

As a model organism with a vast amount of available genetic tools, *A. nidulans* is very well studied with respect to its secondary metabolism. Many efforts have been put into identifying secondary metabolites, and since the genome sequence of *A. nidulans* was released, many approaches have been undertaken in order to investigate the entire secondary metabolism. The genome sequencing has revealed that the *A. nidulans* genome contains 32 putative PKSs, although only a fraction of this number represents the amount of known polyketides from *A. nidulans*.

Preceding genome sequencing of *A. nidulans*, only two genes were linked to a PKS: the conidial pigment YWA1 was linked to the *wA* gene by heterologous expression of the *wA* gene in *A. oryzae* [Watanabe et al., 1998, 1999, Clutterbuck, 1972], and sterigmatocystin was linked to *stcA* (first named *pksST*) by functional inactivation of the *stcA* gene [Yu and Leonard, 1995, Brown et al., 1996].

Following genome sequencing many approaches have been applied for linking genes to PKSs: global and local activation of gene clusters, and co-cultivation with other organisms. Global activation strategies comprised deletion of *sumO* encoding a protein involved in sumoylation, linking asperthecin to the *aptA* PKS encoding gene [Szewczyk et al., 2008], and using chromatin-regulator *cclA* loss-of-function mutants, identifying the monodictyphenone PKS encoding gene *mdpA* [Bok et al., 2009]. Aspyridone was linked to the PKS-NRPS encoded by *adpA* by over-expression of the *adpR* TF by random integration [Bergmann et al., 2007]; hybrid PK-NRP emericellamides were linked to *easB* by randomly deleting six NRPSs [Chiang et al., 2008]; and asperfuranone was linked to the PKSs encoded by *afoE/afoG* by over-expression of the TF, replacing the promoter of the TF with *PalcA* [Chiang et al., 2009]. A creative and interesting approach was applied for linking orsellinic acid and derivatives to *orsA*; co-cultivating *A. nidulans* with 58 different actinomycetes [Schroeckh et al., 2009]. Recently, aspernidine A was identified in a mitogen-activated protein kinase deletion

mutant (*mkpA* Δ) obtained from the genome-wide kinase deletion study [De Souza et al., 2013] and linked to *pkfA* by gene deletion [Yaegashi et al., 2013].

We have used a genome-wide PKS deletion approach in order to identify PKs and link them to their corresponding PKSs (Chapter 4). As a result of this study austinol and dehydroaustinol were linked to the PKS encoded by *ausA*, and by over-expression of *ausA* we found that AusA synthesizes 3,5-dimethylorsellinic acid which serves as a precursor to the austinols [Nielsen et al., 2011]. The biosynthesis towards austinols was subsequently elaborated by Itoh et al. [2012], Lo et al. [2012] and Matsuda et al. [2013].

2.1.1 Systematic over-expression of phosphatases and TF triggers secondary metabolite production in *A. nidulans*

The above mentioned efforts comprise very different approaches and illustrate how creativity is essential in the investigation of fungal secondary metabolism. However, it is striking how difficult it is to trigger secondary metabolite production in a well-characterized model fungus with many available tools. Even by applying all these different approaches, only a third of the potential polyketides are known. This raises the question of whether the availability of the full genome sequence has led us on a wild goose chase for linking imaginary compounds to non-functional genes, which have just been carried down through evolution, or if the majority of these compounds are indeed only expressed under extremely rare and specific conditions.

In order to address this question, I hypothesized that a drastic perturbation of the general metabolism might stress the cell enough to completely reroute the secondary metabolism. Having developed a HTP expression system, I assessed the possibilities of perturbing the metabolism by gene over-expression. Sopko and co-workers had performed a genome-wide over-expression of all genes (>80 %) in *S. cerevisiae*, and found that the cells were sensitive to perturbation of pathways defined by proteins involved in cell-cycle regulation, transcriptional activation, transport and signaling e.g. kinases and phosphatases. Inspired by this result, I decided to over-express all 52 putative TFs from *A. nidulans* chromosome I and all 28 *A. nidulans* phosphatases, under control of the doxycycline inducible O7 promoter (TetON/OFF) to prevent possible lethal phenotypes.

In this study I have investigated these strains only by cultivation on minimal medium, allowing for easy identification of new compounds because *A. nidulans* normally only produces austinols and sterigmatocystin on minimal medium. Indeed, some of the perturbations had a great effect on the metabolite profile. Most strains did not display an obvious change of metabolite profile, and some of the strains seemed to overall down-regulate production of secondary metabolites. Importantly, some of the strains produced many more metabolites than the reference, some of which may be novel compounds. Comparative examination of entire metabolite profiles is a laborious task and should be performed by skilled analytical chemists, why I have only examined a few selected strains in order to generally assess the response to over-expression of these genes.

Thus, examination of six selected strains verified that it is indeed possible to alter the metabolic profile of *A. nidulans* by systematic over-expression of phosphatases and TFs. Dereplication of the six strains revealed

a high production of compounds that are not normally produced on minimal medium e.g. desferritriacetyl-fusigen (NRP), arugosins and shamixanthones (PKs), violaceols (PKs), and asperugin B (PK). Interestingly, I have also identified four unknown compounds which are produced in high amounts. UV spectra indicate that three of them are NRPs, and one may be a heterodimer of a sterol and another compound, a configuration which was only recently reported by Miao and co-workers [Miao et al., 2012]. Based on the fact that the external environment (medium composition, pH, temperature etc.) has a great impact on regulation of secondary metabolism, I hypothesize that cultivation of these strains under different circumstances (e.g. complex media) will provide an even stronger response to the gene over-expression and perhaps result in production of many novel compounds.

Having identified possible novel compounds, it is important to keep in mind that we are studying a perturbed metabolic system. Thus there is a possibility that the new compounds are not naturally produced by the fungus, but are products of cross-talk between different biosynthetic pathways, which have been triggered by abnormal regulation due to over-expression of the genes. Even so, identification of these compounds may help link compounds to genes and learn more about secondary metabolism in general, and even about cross-talk between defined pathways. Additionally, novel compounds may also prove bioactive and be of commercial value, whether or not they are produced naturally.

Although only some of the strains seem to up-regulate production of secondary metabolites, I consider this study a success. I have confirmed that HTP gene expression is extremely fast and non-laborious using the platform described in this study, why it is quick and easy to construct many strains. Random over-expression of genes is not a profound intellectual approach, however, the current knowledge on regulation of secondary metabolism is apparently not sufficient for understanding the complexity of these mechanisms. Using a random approach may, in addition to detection of novel compounds, help identify genes that are involved with local and global regulation mechanisms and shed some light upon the specific reactions employed by the cell. Being able to prepare fungal strains in a HTP manner eases the application of random gene expression, as the need for a high success-rate is less pronounced. "Quick-and-dirty" screening can even be accomplished using the AMA1-based expression platform - also in *Aspergilli*, for which no genetic tools exist.

2.1.2 Assessment of novel genetic tools

In this study I have developed a system for HTP gene expression in *A. nidulans*. The system comprises two different parts: a background-free cloning vector and a background-free integration system for *A. nidulans*. The background-free cloning vector contains the *ccdB* suicide gene, which is excised when the vector is properly digested. This eliminates background from non-digested vector, and results in very high cloning efficiencies. The background-free integration system is based on prior integration of the *wA* gene (or another pigment producer) into a defined integration site, using *pyrG* as a selection marker. Subsequently *your favorite gene* is integrated on top of the prior construct (using a different selection marker), thus looping out *wA* and *pyrG*. Addition of 5-FOA, which is toxic to *pyrG*-expressing cells, will eliminate growth of false

positive transformants as random integration will not excise *pyrG*. In addition, the YWA1 pigment will no longer be produced, and a clear color change will be observed. Using only the color screen for correct integration, it is even possible to visually distinguish between correct transformants and false positives (see Figure 2.1).

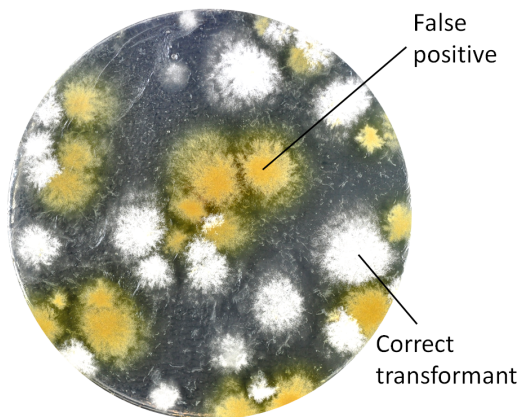


Figure 2.1: Transformation plate using the background-free integration system for *A. nidulans* without addition of 5-FOA to the medium.

In this study I have used the background-free HTP expression platform for over-expression of *A. nidulans* phosphatases and TFs. As expected, strain construction was very fast and efficient. Using the *ccdB* suicide gene in the cloning vector truly eliminates false positives, and in combination with the robust and buffer-tolerant USERTM enzyme it even allows for cloning with unpurified PCR products. The background-free integration system for *A. nidulans* worked very well, and made it easy to identify correct transformants.

Chiang and co-workers recently described an approach where they integrated genes into the *yA* locus of *A. nidulans*, thus resulting in yellow conidia instead of the green conidia produced in the wild-type [Chiang et al., 2013]. Curiously, this is the same approach as I have used in *A. niger* for one of the novel integration sites I have identified. Although this approach works well, it is not very flexible as it only allows for integration into one specific site. As gene expression is highly locus-dependent a library of different available integration sites can be used to choose an appropriate level of expression, preferably in combination with a library of different promoters. Thus I have identified 11 new promoters and five new integration sites for *A. nidulans*. Expression of *lacZ* and *wA*, respectively, confirms that expression levels are very locus-dependent (see Figure 2.2). Discovery of IS5 was a very important result of this study. Providing an extremely high expression level, this site can be used for high production of proteins or small molecules. In fact, with expression from this site (plus a small contribution from IS1), *A. nidulans* is able to direct more than 22 % of the total carbon into production of 6-MSA without significantly affecting the growth rate.

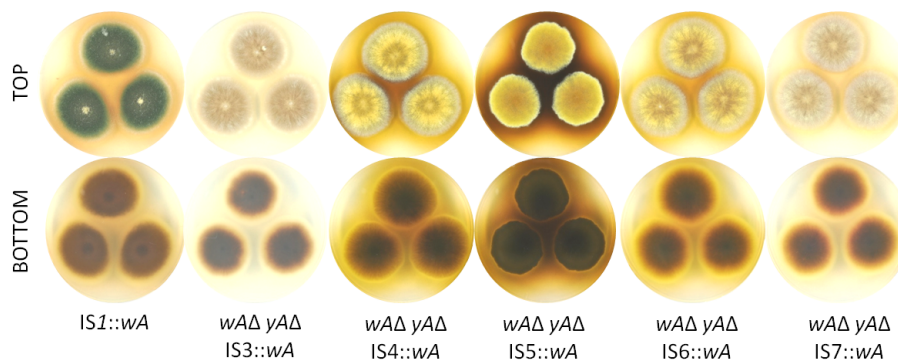


Figure 2.2: Production of YWA1 from six different integration sites in *A. nidulans*.

Even though high expression levels are required for some applications, it is equally important to be able to choose low level gene expression. Low level expression is especially important for metabolic engineering [Pitera et al., 2007], for production of toxic compounds, or for certain specific studies. In this study, integration site IS3 (low expression) was used for testing autoregulation of the (+)-geodin TF, using the reporter gene *lacZ*. If in fact autoregulation did happen, this effect could have been masked by saturation if the protein was highly expressed prior to induction.

2.2 Investigation of secondary metabolism in *A. niger*: a fungus with few genetic tools

When this study was initiated, none of the metabolites in *A. niger* were linked to their corresponding gene (cluster). However, it was known that *A. niger* possessed a gene cluster with strong homology to the fumonisins gene cluster in *G. moniliformis* [Pel et al., 2007, Andersen et al., 2011], and it was shown that some strains of *A. niger* were able to produce fumonisins [Frisvad et al., 2007].

During the past three years there have been a lot of advances within the field of linking genes to their corresponding PK(s) in *A. niger*. The first discovery was made simultaneously by Jørgensen and co-workers [Jørgensen et al., 2011] and Chiang and co-workers [Chiang et al., 2011] when they reported that the FwnA/AlbA PKS is responsible for production of the black conidial pigment as well as production of naphtho- γ -pyrones. Subsequently, Li and co-workers discovered that *adaC* encodes a PKS responsible for synthesis of tetracyclins [Li et al., 2011b]. This PKS was not previously identified in the *A. niger* genome, as the sequence was incomplete. Li and co-workers obtained the full sequence by PCR and sequencing, and reported the domain structure SAT KS AT PT ACP [Li et al., 2011b]. Gil Girol and co-workers identified the PKS responsible for production of bicoumarins (e.g. kotanins and funalenone), encoded by *ktnS* [Gil Girol et al., 2012] and finally, Zabala and co-workers reported the silent azaphilone gene cluster (*aza*) that is somewhat similar to the asperfuranone gene cluster in *A. nidulans* [Zabala et al., 2012]. The otherwise silent *aza* cluster was activated by over-expression of a TF embedded in the cluster by random integration.

Compared to *A. nidulans*, genetic engineering in *A. niger* is more laborious, as only few genetic techniques are available. Elimination of non-homologous end-joining by deletion/disruption of *kusAΔ* (Ku70 homologue) greatly enhances the transformation efficiency [Meyer et al., 2007], and in combination with the use of bipartite (split marker) transformation substrates [Nielsen et al., 2006] it is now possible to perform gene targeting very efficiently. One of the remaining issues with genetic engineering in *A. niger* is the availability of selection markers. Most *A. niger* strains are prototrophic and thus only allow for use of dominant selection markers e.g. *hph* or *bleR* conferring resistance to hygromycin B and phleomycin, respectively. The few available markers are not counter-selectable, thus greatly limiting the number of possible gene targeting events. Furthermore, the lack of a known sexual cycle prevents combination of mutations by sexual crossing. Another Ph.D. student in our research group is working to solve this problem, so I have not dealt with this issue in the present study.

Another disadvantage of genetic engineering in *A. niger* is that there are no defined integration sites for ectopic or heterologous gene expression. It is important to use defined loci for ectopic expression as chromosomal positioning has a large impact on gene expression levels and epigenetic effects [Verdoes et al., 1995, Sousa et al., 1997, Thompson and Gasson, 2001]. I have identified four integration sites for *A. niger* using three different approaches: IS-NIG1 was identified based on gene expression data from a transcriptome analysis, IS-NIG2 disrupts *albA* resulting in white transformants, and IS-NIG3/IS-NIG4 were identified based on homology to the high-expression integration site IS5 in *A. nidulans*. IS-NIG1, 3, and 4 have not yet been thoroughly investigated. Integration into IS-NIG2 greatly facilitates verification of correct transformants, as the color shift from black to white is visibly detectable. Furthermore, disruption of *albA* abolishes production of naphtho- γ -pyrones which constitute the largest group of compounds produced by *A. niger* at all conditions [Nielsen et al., 2009], and provides a clearer background for studying secondary metabolites. Coincidentally, development of this integration site was essential for linking a second copy of the Alba PKS, NapA, to naphtho- γ -pyrone production (Chapter 9).

2.2.1 Heterologous expression links three PKSs to compounds

In this study I have linked three *A. niger* PKSs to known *A. niger* compounds: YanA produces 6-MSA, which is a precursor to yanuthone D, NapA produces YWA1 - the precursor to naphtho- γ -pyrones, and ASPNIDRAFT_44005 produces asperrubrol using cinnamic acid as a starter unit. Furthermore, I have deduced the biosynthetic pathway towards yanuthone D using a combination of individual cluster gene deletions in *A. niger*, heterologous expression of cluster genes in *A. nidulans* and feeding experiments with labeled and unlabeled metabolites.

The overall approach of investigating the *A. niger* polyketome by heterologous expression of PKSs in *A. nidulans* resulted in only a few success stories. From 37 PKSs (and hybrid NRPS/PKSs), only two compounds (6-MSA and YWA1) were identified by chemical analysis. One PKS (NapA) provided an altered phenotype, but it was far from trivial to deduce the actual PK that was produced (YWA1). A combination

of gene deletions and over-expression in *A. niger* resulted in many pieces of an overall puzzle, which was very complex and difficult to solve. In fact, and as mentioned above, the YWA1 product of NapA was only discovered because I coincidentally used the *albA* locus as an integration site for over-expression of *napA*, resulting in a white *A. niger* strain that still produced naphtho- γ -pyrones. AlbA and NapA apparently both produce YWA1, which is the precursor to naphtho- γ -pyrones and black melanin in *A. niger*. However, as *napA* seems to be silent under most circumstances, this gene could not have been linked to its product by traditional gene deletion methods. Interestingly, *napA* is part of a typical gene cluster encoding all necessary activities for conversion of YWA1 into naphtho- γ -pyrones, whereas AlbA does not appear to be part of a gene cluster. Due to the similarities of the two proteins, they must have originated via divergent evolution. I propose that one may have formed from the other by gene duplication, or that one (or both) have emerged as a result of horizontal gene transfer. Unfortunately, I did not have time to finish this study before submitting this thesis. I am still to construct and analyze two strains before submitting this study for publication: a double deletion of *albA* and *napA* to examine if NapA is responsible for the fawn, and not entirely white, conidia of the *albA* Δ strain, and a *nap* cluster deletion strain, which I expect will verify that without the presence of the *nap* cluster, *A. niger* is still able to produce black conidia, but it is dependent on the *nap* cluster and is no longer able to produce any naphtho- γ -pyrones.

Asperrubrol is a very good example of a PK that would not be linked to a PKS gene only by heterologous expression approaches. In this case, heterologous expression of the gene in *A. nidulans* did not provide any new products. The PKS did not contain any release domains, but there was a gene encoding a discrete thioesterase located downstream of the PKS, which could serve as the release mechanism. Expression of both the PKS and the thioesterase in *A. nidulans* did, however, still not provide any products. Intrigued by this result, I deleted the PKS gene in *A. niger*, and examination of the metabolite profile revealed that asperrubrol was no longer produced by the deletion strain. Investigation of a proposed gene cluster suggested that cinnamic acid might be a starter unit to the PKS. Indeed, lack of a suitable starter unit explains why heterologous expression of the PKS did not result in a product. Heterologous expression of a cluster gene did in fact produce cinnamic acid, substantiating this hypothesis. Thus, this example further illustrates the benefit of combining classical gene deletion approaches with heterologous expression in a suitable host.

2.2.2 Combining several approaches reveals the biosynthesis of yanuthone D in *A. niger*

Although heterologous expression of *yanA* in *A. nidulans* identified 6-MSA as the product, it was not obvious that yanuthone D was the end product. Thus *yanA* was only linked to yanuthone D by gene deletion in *A. niger*. Reversely, if only gene deletion had been carried out the surprising discovery that 6-MSA is a precursor to yanuthone D might not have been made. Even individual deletion of cluster genes or over-expression of *yanA* in *A. niger* did not result in accumulation of 6-MSA, why heterologous expression of *yanA* in *A. nidulans* indeed proved valuable for deduction of the yanuthone D biosynthetic pathway. In this particular study, I

have combined a range of different approaches for investigation of the biosynthetic pathway. Besides gene deletions in *A. niger* and heterologous expression of *yanA* in *A. nidulans*, I have used a novel plasmid-based expression platform for "quick and dirty" screening of cluster genes in *A. nidulans*. This approach provided the first tailoring enzyme to act on 6-MSA in the yanuthone D pathway. The first part of the yanuthone D pathway was not accounted for by gene deletions, why this tool helped investigation of the biosynthesis. In addition, feeding with fully labeled 6-MSA was carried out, and confirmed that 6-MSA is indeed incorporated into yanuthone D. Using the newly developed integration sites I constructed a strain expressing *yanA* from IS1 and IS5, resulting in a very high production of 6-MSA. In fact we were able to produce titers of 1.81 g/l (0.227g/g glucose) fully labeled 98.7 % isotopic pure 6-MSA from fully ^{13}C labeled glucose.

The complete yanuthone D study brilliantly illustrates the benefits of combining different approaches, and how tool development and optimization can be essential for studying complex biological systems. Development of new integration sites may seem trivial and un-inventive, but if IS5 had not been discovered, production of fully labeled 6-MSA would probably have been too expensive. Similarly, development of integration sites in *A. niger* coincidentally solved the complex puzzle of YWA1 production from AlbA and NapA.

2.3 Investigation of secondary metabolism in fungi with no available genetic tools

As briefly mentioned in the introduction, 64 % of all new drugs are natural products or natural product mimics [Newman and Cragg, 2012]. This number indeed suggests that there is a huge demand for discovery of many novel bioactive compounds, if human kind is to continue medicinal development. Plants and filamentous fungi are great sources of natural products, and it is estimated that there are more than 1,400,000 species of filamentous fungi [Petersen, 2003]. Secondary metabolite production can vary greatly between species of the same genus e.g. *A. nidulans* and *A. niger* barely share any common secondary metabolites. Mycologists and natural product chemists are doing an amazing job investigating all sorts of exotic fungi, and they often succeed in discovering novel compounds using traditional mycological methods such as varying medium composition, pH, temperature ect. However, in the light of recent genome mining studies it has become clear that we only see the tip of the iceberg when studying fungi by traditional methods in the laboratory. Even for *A. nidulans*, whose secondary metabolism is extremely well-studied by traditional and several different genetic approaches, only a fraction of the potential compounds are known.

Thus there is also a need for discovery of secondary metabolites that are not produced under standard laboratory conditions, but this is a very difficult problem to solve. For a thorough investigation of the secondary metabolism in a given fungus, it is necessary to have a system for genetic engineering. To set up such a system can be extremely tedious, and may be different for each species. Even the most basic techniques e.g. transformation of DNA into the cell, may be difficult to implement. The lack of development of genetic

techniques for new fungi is in contrast to the steadily increasing number of whole genome sequences, which allows for easy identification of candidate genes.

Thus, in order to investigate secondary metabolism in a fungus with no available genetic techniques there are basically two good options: heterologous expression of genes in a host with well-characterized genetic techniques or expression of genes (or RNA) in the native fungus from an AMA1-based plasmid. Both approaches have advantages and disadvantages, but in combination they may serve as a powerful tool.

We have successfully transferred the entire (+)-geodin cluster from *A. terreus* into the well-characterized host *A. nidulans*. The cluster contained a TF, which was constitutively expressed in *A. nidulans* for positive regulation of the entire gene cluster. After transfer of the cluster, it was possible to apply the *A. nidulans* extensive genetic toolbox for investigation of individual genes in the (+)-geodin pathway.

The other approach I have used involved a combination of heterologous expression and expression in the native host from AMA1-based plasmids. In this case I wanted to examine a genome-sequenced *Aspergillus*, but for which no genetic tools existed. Only a few compounds had been reported for the industrial enzyme producer *A. aculeatus* why this fungus posed an almost clean slate for studying secondary metabolism. I first examined the fully sequenced genome of *A. aculeatus* in a search for interesting candidates, and I have included the results of this investigation in Appendix H. Interestingly, the genome of *A. aculeatus* contains 57 putative synthases, of which 24 are PKSs. Phylogenetic analyses revealed that one of the PKSs was very similar to fungal 6-MSA synthases e.g. *yanA* from *A. niger*, suggesting that *A. aculeatus* might produce a PK that uses 6-MSA as a precursor. As it is known that 6-MSA can be converted into toxic compounds e.g. patulin, it was relevant to examine if this industrial fungus is in fact able to produce toxic compounds. First, I verified that the putative 6-MSA synthase did produce 6-MSA by heterologous expression in *A. nidulans*. As this was the case, the next logical step was to delete the gene in *A. aculeatus* in order to link the PKS to a compound. In spite of several attempts and different markers, I was not able to delete the 6-MSA synthase gene in *A. aculeatus*. Instead I over-expressed the 6-MSA gene from an AMA1-based plasmid in *A. aculeatus*, hoping to link the gene to a compound by up-regulating production of this compound. This approach resulted in production of 6-MSA, but the remainder of the metabolite profile remained unchanged. This result suggested that if 6-MSA was a precursor to another compound, the appertaining gene cluster might be silent under laboratory conditions. Therefore, I examined a possible gene cluster for a transcription factor, and I located a putative TF just downstream of the 6-MSA synthase. Over-expression of the TF in *A. aculeatus* indeed activated a biosynthetic pathway, resulting in production of five new compounds, of which two have been structure elucidated by NMR spectroscopy. Feeding with fully ^{13}C labeled 6-MSA confirmed that 6-MSA was incorporated into three of the compounds, which have not been structure elucidated at present time.

As briefly mentioned above one option for investigating secondary metabolism in fungi with no genetic tools is by expression of RNA. Expression of antisense RNA or RNA interference (RNAi) is a way of silencing a gene without having to delete the gene. This is a great advantage in fungi where it is not possible to

do direct genetic alterations, as the RNA can simply be expressed from AMA1-based plasmids. The use of antisense RNA has been successfully employed for filamentous fungi e.g. for down-regulation of protease production in industrially relevant strains of *A. oryzae* and *A. awamori* [Zheng et al., 1998, Moralejo et al., 2002]. As silencing with antisense RNA is easy, but not very efficient, is it also possible to silence a gene by RNAi. Silencing by RNAi is more efficient [Nakayashiki et al., 2005], but it is also more challenging to design an optimal RNAi. Furthermore, studies have shown that RNAi constructs are unstable and lost throughout cultivation of *N. crassa* and *A. fumigatus* [Goldoni et al., 2004, Henry et al., 2007]. However, for studying secondary metabolism, strains are usually not subjected to prolonged cultivation and the effect should last long enough to detect the change in the metabolite profile. Thus I propose that a combination of heterologous expression, gene over-expression from AMA1-based plasmids, and silencing by RNA could pose a powerful tool for investigating secondary metabolism in genome-sequenced fungi, for which no genetic tools exist.

2.3.1 *A. nidulans* as a host for heterologous gene expression

During this study *A. nidulans* was used as a host for heterologous expression of genes from *A. niger*, *A. aculeatus*, *A. carbonarius*, and *A. terreus*. For *A. niger* genes, *A. nidulans* was used for individual expression of 37 putative PKSs, as well as expression of selected cluster genes to illuminate biosyntheses of yanuthone D (Chapter 10) and asperrubrol (Chapter 8). Additionally, *A. nidulans* was used for verification of a 6-MSA synthase from *A. aculeatus*, and a YWA1 synthase from *A. carbonarius*. Last, *A. nidulans* was used as a host for transfer of the entire (+)-geodin cluster from *A. terreus* and subsequent examination thereof.

Generally, the abundance of genetic tools for *A. nidulans*, compared to other Aspergilli, makes *A. nidulans* a suitable host for heterologous gene expression. Development of NHEJ deficient strains significantly increases the transformation efficiencies, and in combination with the use of bipartite/split markers, screening for correct transformants is no longer tedious. Application of the background-free integration system, developed in the study, further increases the ease of gene expression in this fungus. Identification of the five novel integration sites also allows for simultaneous integration of more than one gene in a single reaction, and enables individual expression of several cluster genes, regulated by defined promoters.

Heterologous expression of 37 putative *A. niger* PKSs resulted in a very low success rate, as only three PKSs were linked to a compound. However, the reason for the low success rate is most likely not due to *A. nidulans* being the host. I assess that the problems were primarily due to incorrect prediction of the PKS encoding genes, lack of specific starter units, or lack of a release mechanism. Thus, over-expression of the genes in *A. niger* instead of *A. nidulans* might have provided a compound for the PKSs lacking a starter unit or a release mechanism. Development of four new integration sites in *A. niger* enables this possibility. Using *AlbA* as an integration site would allow for fast verification of correct transformants and would even make identification of new compounds easier, as there would be no production of naphtho- γ -pyrones.

The approach of transferring the entire (+)-geodin cluster to *A. nidulans* was one way of overcoming the obstacle of performing genetic engineering in *A. terreus*. The approach was successful, however, not very

efficient. Even by constitutive expression of the TF, only very small amounts of (+)-geodin were detected. Thus the approach could be optimized by individual expression of the cluster genes, each controlled by a defined promoter. Such an approach would be more time-consuming for preparation of the strains, and one would have to make sure that ORF of the genes are predicted correctly. Hence, this approach would also serve as a verification of correct gene prediction and aid in development of future algorithms for automatic gene prediction in filamentous fungi.

One of the reasons why *A. nidulans* is a good choice for heterologous expression of biosynthetic genes from e.g. *A. niger*, *A. aculeatus*, *A. terreus*, and *A. carbonarius* is that *A. nidulans* only share a few common secondary metabolites. Although all belong to the genus *Aspergillus*, they belong to different clades [Varga et al., 2003] and are relatively distantly related. The distant relation may have allowed for adaptive evolution and horizontal gene transfer within a defined ecological niche, and would explain why the metabolite profiles of these species are so different. Thus, examination of secondary metabolism of closely related species (e.g. *A. ustus* or *A. versicolor*) by heterologous expression in *A. nidulans*, may not prove as easy.

A. nidulans is most often regarded a model organism for studying cell biology in the lab, however, the fact that *A. nidulans* is able to produce immense amounts of 6-MSA and YWA1 without being significantly impaired, suggests that *A. nidulans* might also be a good cell factory. *A. nidulans* does not naturally produce enzymes or compounds of interest, and probably for that reason it was never considered an industrial workhorse. However, it also does not produce vast amounts of by-products (e.g. citric acid), and it is very well characterized. Thus, it may be possible that *A. nidulans* may have a future as a cell factory.

2.4 Concluding remarks and future perspectives

In this study I have used a wide range of different genetic approaches for investigation of secondary metabolism in Aspergilli including gene over-expression, heterologous gene expression, gene deletion, and site-directed mutagenesis. For this purpose I have developed a series of new genetic tools for gene expression, including a background-free HTP expression platform for *A. nidulans* and an AMA1-based expression platform for fungi with no available genetic tools.

Genome-wide deletion of PKSs in *A. nidulans* identified the PKS responsible for production of austinols in *A. nidulans*, AusA. Furthermore, HTP over-expression of *A. nidulans* phosphatases and transcription factors resulted in altered metabolite profiles, and a brief inspection revealed four possible novel compounds.

The combination of genetic engineering and feeding experiments has helped reveal the PK product of four *A. niger* PKSs. Furthermore, the entire biosynthesis towards yanuthone D in *A. niger* was deduced, and it was discovered that *A. niger* contains two different PKSs that produce YWA1, the precursor to black melanin and naphtho- γ -pyrones. The PKS responsible for production of asperribol was identified using classical deletion methods, and it was shown that this PKS requires a cinnamic acid starter unit. Interestingly, two compounds that were previously classified as terpenoids (yanuthone D and asperribol),

proved to be PK-derived.

I have shown that *A. aculeatus* contains a 6-MSA synthase, which defines a biosynthetic cluster. Over-expression of a TF embedded in the cluster activated the biosynthetic pathway and revealed five novel compounds. Feeding studies verified that 6-MSA was incorporated into three compounds, whose structures have not yet been solved.

In conclusion, I have successfully elucidated many parts of fungal secondary metabolism by combining genetic, mycological and chemical approaches. It is clear that fungal secondary metabolism is very complex, why I predict that many biosynthetic clusters will not be identified merely by deleting or over-expressing genes. Based on this experience, I will advice everyone working in the field of secondary metabolism to join forces to maximize the outcome of these studies. After all, we all share a common goal of making the world a better place.

Chapter 3

Development of novel tools for genetic engineering in Aspergilli

During this study, I have developed new methods for simplifying HTP strain construction, and I have expanded the current toolbox for performing genetic engineering in a simple and flexible manner. These tools and methods have been used sporadically in the following chapters, why I have chosen to briefly introduce the principles of the new tools and application thereof.

In this chapter I will describe the tools that I have developed: new promoters from *A. nidulans* and new integration sites for *A. nidulans* and *A. niger*. Furthermore, I will describe the methods that I developed in order to facilitate HTP strain construction in general: development of a plasmid-based expression platform for Aspergilli, implementation of the *ccdB*-based background free UE-compatible cloning vector, and development and implementation of a background-free transformation system for gene expression in *A. nidulans*.

3.1 Results and discussion

3.1.1 Expanding the library of promoters for *A. nidulans*

One thing that is important for strain construction is the availability of usable promoters for a suitable expression level of *your favorite gene* (YFG). Until now, only a limited number of promoters have been reported for use in Aspergilli e.g. the strong, constitutive promoters *PgpdA* from *A. nidulans* and the glucoamylase promoter *PglaA* from *A. niger* [Ganzlin and Rinas, 2008]. Very recently, six new constitutive promoters with varying strengths were characterized for use in *A. niger* [Blumhoff et al., 2013]. Two inducible promoters have also been described for use in Aspergilli: the alcohol-inducible *PalcA* from *A. nidulans* and the sucrose-inducible *sucA* from *A. niger*. The latter was only recently described [Zuccaro et al., 2008, Roth and Dersch, 2010] why the potential of this inducible promoter is not yet known. Additionally, our research group has recently implemented use of the tunable O7 promoter (TetON/OFF) in *A. nidulans*, which was

adapted to *A. niger* by Meyer and co-workers [Meyer et al., 2011].

Besides the need for strong, constitutively expressed promoters, there is also a general need for new constitutively expressed promoters that do not necessarily have to provide high expression. This need has emerged because it has become possible to express entire gene clusters in a heterologous host, with the aim of uncovering unknown biosyntheses (Chapter 11) [Chiang et al., 2013]. If the genes are individually expressed, it is an advantage to employ different promoters, as this will decrease the risk of spontaneous direct repeat loop-out events. Especially in the field of secondary metabolite expression, compounds that are toxic to the cell may induce a stress response, which could potentially increase the rate of recombination in an effort to get rid of the genes producing the toxic compound.

Eukaryotic promoters are complex and diverse, why they can be very difficult to identify and characterize. Regulatory elements may be present several kilobases from the transcriptional start, and these are very difficult, if not impossible, to predict. The characterized *A. nidulans* promoters *PgpdA* and *PalcA* are 2.3 kb and 800 bp, respectively, and the *A. niger* glucoamylase promoter *PglaA* is 2 kb long. Note that these are the promoter lengths that are applied by molecular biologists; the entire promoter may contain regulatory elements elsewhere which have not been identified. Generally, a short promoter is preferred because it eases PCR reactions and keeps vectors at a minimum size (many fungal expression vectors are >13 kb without gene inserted), but it also decreases the risk of spontaneous recombination within the cell.

In this study I have tested 11 new promoters fra *A. nidulans*, based on micro-array data (kindly provided by Mikael R. Andersen). With only one exception, I used 1 kb as a cut-off value. 8 of the promoters provided good expression of genes, with varying expression levels, see Figure 3.1. The results are summarized in Table 3.1. One of the promoters with strong is a short (200 bp) strong bidirectional promoter (P_{AN4802}/P_{AN4803}).

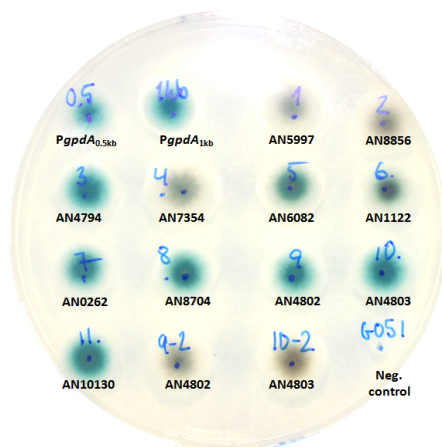


Figure 3.1: Qualitative test of new promoters on X-gal plate. The amount of blue color is an indirect measure of *lacZ* expression.

Table 3.1: *A. nidulans* promoters analyzed in this study. AVG exp = average expression level in the A4 wiild-type strain on all media. A4 glc = expression levels in A4 on glucose, A4 gly = expression levels in A4 on glycerol, n.d. = not determined in this study, *Qualified based on β -galactosidase assay (Figure 3.1).

Locus	AVG exp	A4 glc	A4 gly	Length (kb)	Expression at IS1*
AN0262	13.95	14.01	13.95	1	Strong
AN1122	14.06	13.99	14.06	1	Intermediate
AN4794	14.15	14.18	14.12	1	Strong
AN4802	13.96	13.91	13.94	0.2	Strong
AN4803	13.84	13.80	13.85	0.2	Strong
AN5997	14.21	14.32	14.14	1	Weak
AN6082	14.09	14.12	14.10	1	Intermediate
AN7354	14.13	14.17	14.10	1	Weak
AN8704	13.84	13.82	13.84	1	Strong
AN8856	13.99	14.02	13.97	1	Weak
AN10130	14.38	14.42	14.43	1	Strong
<i>gpdA</i>	14.33	14.51	14.37	2.3	Strong
<i>gpdA</i>	14.33	14.51	14.37	1	Strong
<i>gpdA</i>	14.33	14.51	14.37	0.5	Strong
<i>alcA</i>	10.41	9.33	14.28	0.8	n.d.

Six of the promoters provided strong expression, comparable to expression levels obtained with *PgpdA*. All of these promoters belonged to ribosomal proteins (AN0262, AN4794, AN4802, AN4803, AN8704, and AN10130). A single promoter belonging to a ribosomal protein, AN6082, only provided intermediate expression levels based on the quantitative X-gal assay.

AN5997 is involved in rRNA export, and AN8856 has a role in rRNA subunit assembly. Both of them displayed weak expression. The two remaining promoters are from genes that encode proteins with no putative function (AN1122 and AN7354), and displayed intermediate and weak expression, respectively.

Based on these data, it is clear that promoters that regulate expression of ribosomal proteins, are highly expressed under the circumstances used in this study. This study was carried out on minimal medium (MM), which is also the cultivation medium most frequently used in our laboratory, why promoters of genes encoding ribosomal proteins are good targets for development of new constitutively expressed promoters for use in the laboratory. For other applications, promoters should be tested accordingly.

The availability of these new promoters allow for gene expression of several individual genes, regulated by different constitutive promoters. This is a clear advantage for reconstitution of biosynthetic clusters, however, I have not yet had an opportunity to use these promoters for such application.

3.1.2 Development of five new integration sites in *Aspergillus nidulans*

For heterologous expression of genes, one may choose to either integrate the gene randomly or targeted to a specific locus within the genome. Both approaches are commonly used, and each has its advantages and disadvantages. Random integration is easy because it basically only requires a substrate containing the gene and a selection marker, and it is efficient for high expression as many copies may be integrated during a single transformation event. In contrast, targeted integration requires a substrate with at least the gene, the selection marker, and two flanking *targeting sequences*. Usually, only one copy is integrated during transformation, however it is possible (but not trivial) to do two or more targeted integrations, especially when using multiple selection markers. The great benefit, however, of targeted integration is that expression levels are constant for all genes expressed from that locus. Thus gene expression from a defined locus is highly suitable for screening purposes e.g. when comparing promoter strengths or enzyme activities. Moreover, the risk of disrupting genes will be prevented.

Many studies have been conducted in order to investigate the epigenetic effects of gene insertion, and it is clear that chromosomal positioning has a large impact on gene expression levels in all organisms e.g. gram negative and gram positive bacteria, yeast, filamentous fungi, insects, and humans [Sousa et al., 1997, Thompson and Gasson, 2001, Yamane et al., 1998, Verdoes et al., 1995, Markstein et al., 2008, Gierman et al., 2007]. Thus it is important to have characterized sites for ectopic integration, which can be used for standardized gene expression.

The work I have done on development of new integration vectors for *A. nidulans* is based on the integration system that was developed by Hansen and co-workers [Hansen et al., 2011]. The system is based on UE cloning, and allows for integration into a specific locus termed Integration Site 1 (IS1). This system is highly efficient, but limited because it is only possible to integrate at one site, although it is possible to re-use the *pyrG* marker by counter-selection. Therefore, I wanted to expand the number of integration sites so it would be possible to integrate several genes e.g. for reconstruction of a biosynthetic pathway, either by successive transformation with the counter-selectable *pyrG* marker or by simultaneous transformation using different markers.

The procedure for locating new integration sites (ISs) was straight forward: I located highly expressed regions in the genome (similar to identification of new promoters, see section 3.1.1), and located a potential IS within this region. The potential IS should fulfill the two following criteria: it should be located between two highly expressed genes, and there should be a distance of >1 kb to either of the two ORFs, not to disrupt any surrounding gene or regulation thereof. Based on these criteria five new ISs were tested, using *lacZ* and *wA* as an indirect indicator for expression levels, respectively. An overview of the ISs is shown in table 3.2.

IS	Adjacent gene	Location	AVG exp.*	Distance (nt)	Protein function
1	AN6638 (US)	ChrI	4.35	27	DNA methylation, DmtA [Lee et al., 2008]
	AN6639 (DS)	ChrI	11.2	473	2-methylcitrate dehydratase, McdB
Note: Described in [Hansen et al., 2011]. AN6638 transcript may be disrupted by insertion.					
2	AN4072 (US)	ChrII	9.76	193	Putative histone deacetylase (CADRE)
	AN4073 (DS)	ChrII	13.6	156	Putative 40S ribosomal protein S12 (CADRE)
3	AN4770 (US)	ChrIII	9.80	1,428	PAPS reductase
	AN4769 (DS)	ChrIII	10.8	1,579	ATP sulfurylase
Note: This IS is characterized in chapter 11.					
4	AN4252 (US)	ChrII	12.3	1,017	Unknown
	AN4251 (DS)	ChrII	10.4	1,110	Putative 40S ribosomal protein S2 (CADRE)
5	AN7554 (US)	ChrIV	9.89	1,886	Putative oxidoreductase (CADRE)
	AN7553 (DS)	ChrIV	6.55	2,444	HLH transcription factor, DevR
Note: Very high expression from this site.					
6	AN3180 (US)	ChrVI	9.39	942	Unknown
	AN3179 (DS)	ChrVI	10.0	2,004	Unknown
7	AN8283 (US)	ChrII	12.1	2,259	Putative CYP450, Cyp51B [Kelly et al., 2009]
	AN8282 (DS)	ChrII	7.69	2,287	WD repeat protein (CADRE)

Table 3.2: Insertion sites for *A. nidulans*. *Average gene expression in A4 with D-glucose as carbon source.

For each of the five ISs I constructed expression vectors with new targeting sequences, using *argB* and *pyrG* (with and without direct repeats) as markers, respectively. All vectors contained an AsiSI/Nb.BtsI UE cassette, exactly as described in Section 3.1.5. For each set of vectors, I inserted *lacZ* and *wA*, respectively, and transformed the construct into *A. nidulans*. *ISX::lacZ* transformants (X being the integration site number) were plated on MM plates containing 5-bromo-4-chloro-3-indolyl- β -D-galactopyranoside (X-gal) for qualitative assessment of expression levels (Figure 3.2).

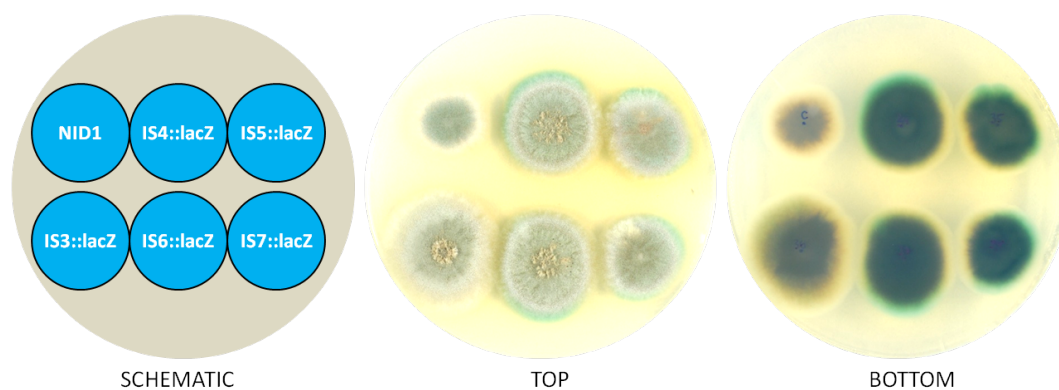


Figure 3.2: Test of *A. nidulans* integration sites using *lacZ* reporter gene.

It was clear that all strains produced a blue pigment compared to the reference, however, it appeared that all ISs had approximately the same productivity except for IS3, which apparently produced less blue pigment than the rest. This was in contrast to the qualitative result obtained with the *wA* expressing strains, which are depicted in Figure 3.3. In this case there was a clear difference in the production of YWA1: IS7 produced approximately the same amount as IS1; IS4 and IS6 had a high production; and IS5 had extremely high production without affecting the growth rate much. Additionally, it was confirmed that IS3 had low production of YWA1.

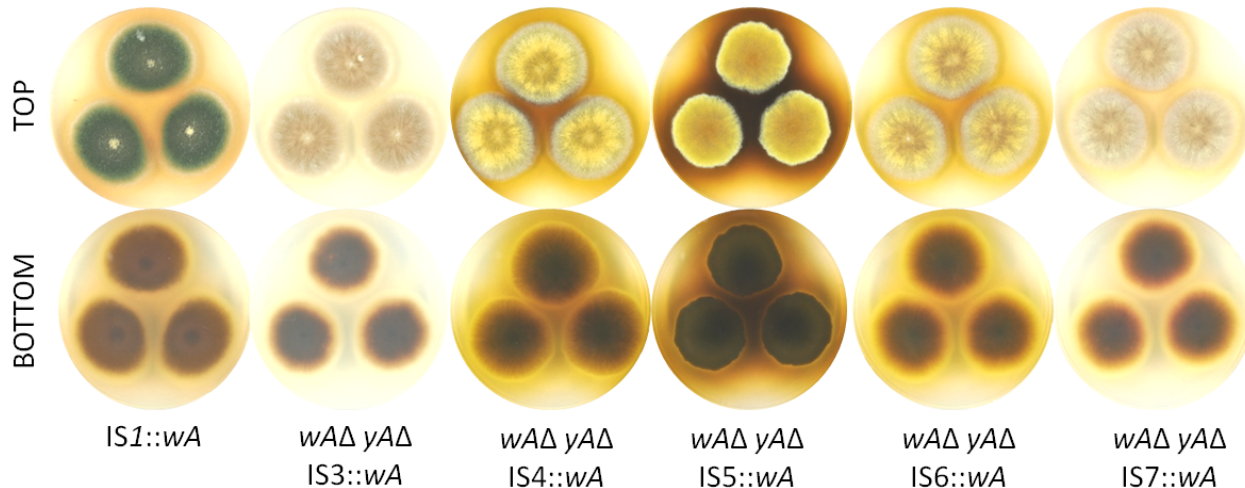


Figure 3.3: Test of *A. nidulans* integration sites using *wA* reporter gene.

IS3 has been used for investigating aspects of the geodin biosynthetic pathway in *A. terreus*, by heterologous expression of cluster genes in *nidulans* (Chapter 11). IS5, providing extremely high production of PKs, has been used for production of fully ^{13}C labeled 6-methylsalicylic acid (6-MSA) for use in labeling experiments (see Chapter 10). In fact, expression from this site combined with expression from IS1 resulted in a titer of 1.81 g/l 98.7 % isotopic pure 6-MSA (this study, data not shown).

3.1.3 Development of four integration sites in *Aspergillus niger*

For productivity purposes random gene insertion is a strong approach, as many copies can be integrated at once, and the resulting strains are screened for productivity and growth. However, in some cases this approach is not optimal: there is a risk that a random gene insertion disrupts important genes and chromosomal positioning also have a large effect on gene expression levels, as shown in section 3.1.2. These "side-effects" are especially important when studying cell biology where gene disruptions might result in a completely different phenotype, and when comparing activities of different genes it is preferable to have identical gene expression levels. For identification of insertion site IS-NIG1, I used the same approach as for *A. nidulans* (see section 3.1.2). For identification of the remaining three insertion sites, I used different approaches, which will be described below. Table 3.3 summarizes data for the four integration sites.

IS	Adjacent gene	AVG exp.*	Distance (nt)	Predicted protein function
NIG1	185605 (US)	9.79	1,645	Mitochondrial 18 KDa protein (CDD)
	56528 (DS)	7.99	473	Histidinol phosphate phosphatase (CDD)
NIG2 (<i>albA</i>)	<i>albA</i> (US)	6.79**	0	YWA1 PKS
	<i>albA</i> (DS)	6.79**	0	YWA1 PKS
Note: disrupts <i>albA</i> gene, resulting in white conidia.				
NIG3	56332 (US)	8.70	4,384	Oxygenase (CDD)
	40564 (DS)	3.79	3,419	No conservation
Note: 56332 is a homologue of AN7554 (<i>A. nidulans</i>)				
NIG4	40564 (US)	3.79	1,723	No conservation
	181931 (DS)	7.64	2,193	HLH transcription factor (CDD)
Note: 181931 is a homologue of AN7553 (<i>A. nidulans</i>)				

Table 3.3: Insertion sites for *A. nidulans*. *Average gene expression in ATCC1015 from liquid culture.

**Secondary metabolism is highly down-regulated in liquid cultivations.

Identification and test of IS2, replacing *albA*/*fwnA* conidial pigment gene

Based on the test of new integration sites for *A. nidulans* (Section 3.1.2) it was clear that the approach of locating highly expressed sites for integration by merely looking at expression levels of surrounding genes, was not consistently successful. Therefore, I wondered if there was another, perhaps more clever, way of designing an integration site for *A. niger*. Screening of *A. niger* transformants is cumbersome because diagnostic (spore) PCR rarely works. As I have noted this problem for several black Aspergilli (*A. niger*, *A. aculeatus* and *A. carbonarius*), but not for the green *A. nidulans*, I assume that the black pigment inhibits the PCR reaction whereas the green does not, at least to the same extent. Based on this hypothesis, it was

clear that an integration site that would abolish production of the black conidial pigment by simultaneous integration of *your favorite gene* (YFG), would be ideal. Moreover, a switch from black to white colonies would make screening for correct transformants very efficient. A second advantage of this approach is that it is not necessary to do the transformations in a *kusA* Δ background, and the *pyrG* marker would be available as a selectable and counter-selectable marker for multiple transformations.

Last, this approach would yield yet another advantage when studying secondary metabolism: *albA* is responsible for production of naphtho- γ -pyrones (e.g. the aurasperones), which constitute the majority of the *A. niger* secondary metabolite pool under all tested conditions [Nielsen et al., 2009]. Hence, the absence of these compounds would make it easier to study secondary metabolites which do not belong to the naphtho- γ -pyrone family. In *A. nidulans* the homologue of *albA*, *wA*, is only expressed in conidia. Thus ectopic gene integration in this site in *A. nidulans*, even if accompanied by a constitutively expressed promoter, may be silenced in other tissue than conidia and may not be suited for studying cell biology. However, very recently, Chiang and co-workers successfully applied this approach to study biosynthesis of secondary metabolites from *A. terreus* [Chiang et al., 2013]. As *albA* in *A. niger* is also responsible for production of naphtho- γ -pyrones, which are detected under all tested conditions, I assume that this gene is not silenced under any known conditions, and that it is expressed in high amounts.

In this case, I could not use *wA* as an indirect marker for gene expression as this gene is the homologue of the gene I was trying to delete. Thus transformants would produce black spores, and the screening system was of no use. Instead I integrated the gene encoding red fluorescent protein (RFP), *rfp*, and the resulting transformants did indeed produce white conidia (Figure 3.4).

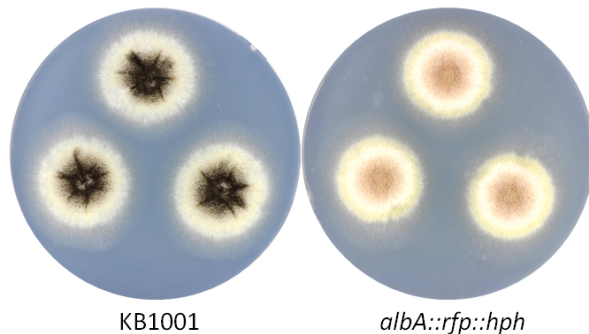


Figure 3.4: Test of *A. niger* integration site 2 (IS-NIG2). Insertion of *rfp* into IS-NIG2 (*albA*) abolishes production of black conidia. KB1001 = reference strain.

Although I did not have time to characterize this integration site any further, it seems it has the potential of being a suitable integration site for HTP strain construction. I have used this integration site, as well as IS-NIG1, for unveiling details of the naphtho- γ -pyrone biosynthesis in *A. niger*, which is presented in Chapter 9.

Identification of IS-NIG3 and IS-NIG4, based on results from *A. nidulans*

Identification of the IS-NIG3 and IS-NIG4 insertion sites were based on the results obtained from the new ISs in *A. nidulans* (see Section 3.1.2). In *A. nidulans* one of the ISs, IS5, had turned out to be a "super-site" for ectopic expression. Therefore I searched for, and found, homologues of the adjacent genes in *A. niger*. ASPNIDRAFT_56332 (56332, AN7554 homologue) and ASPNIDRAFT_181931 (181931, AN7553 homologue) were located closely, however, there was one gene, ASPNIDRAFT_40564 (40564), in between. A BLAST search revealed that the 40564 gene is not conserved among Aspergilli, and expression data indicate that it is very poorly expressed. As I did not know the significance of the 40564 gene, if any, I decided to insert the *wA* gene on the left (IS-NIG3) and on the right (IS-NIG4) flank of 40564, without disrupting any genes.

The resulting transformants did indeed produce high amounts of yellow pigment, however, integration into IS-NIG4 yielded more pigment than integration into IS-NIG3. Furthermore, integration of *wA* into IS-NIG4 seemed to affect growth of the fungus compared to integration into IS-NIG3 (see Figure 3.5). This is an interesting observation as the integration sites are so closely positioned, and only separated by the 40564 gene. Unfortunately, I did not have time to pursue this any further, and these two new ISs have not yet been used any further. However, the strains did not display any devastating morphologies, why both sites are potentially suitable for gene expression.

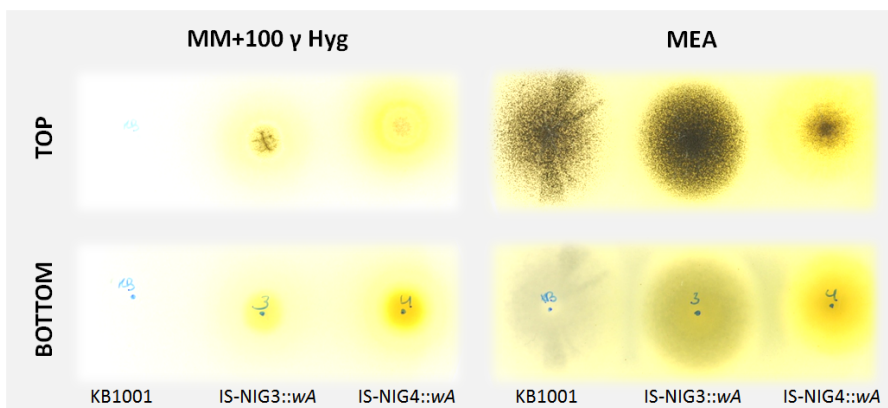


Figure 3.5: Test of *A. niger* integration sites 3 and 4. Insertion of *rfp* into IS-NIG2 (*albA*) abolishes production of black conidia.

3.1.4 Development of a non-integrative (AMA1-based) expression platform for *Aspergillus* Spp.

In addition to the newly developed integration sites, I have constructed an expression cassette within an AMA1-containing non-integrative vector. The AMA1 sequence, which comprise two long perfect inverted repeats separated by a unique spacer, allows for autonomous plasmid replication in *A. nidulans*, *A. niger*, and *A. oryzae* [Gems et al., 1991, Aleksenko et al., 1996].

Similar to existing integration vectors, the AMA1 expression vector contains the *PgpdA*_{0.5kb} promoter and the *TtrpC* terminator. I constructed three vectors with different markers: *argB* (pDHX1), *pyrG* (pDHX2) and *hph* (pDHX3) (see Figure 3.6). The latter, containing the hygromycin phosphatase *hph* marker, is independent of auxotrophies and can be used for gene expression in wild-type *Aspergilli*, and maybe even filamentous fungi belonging to other genera.

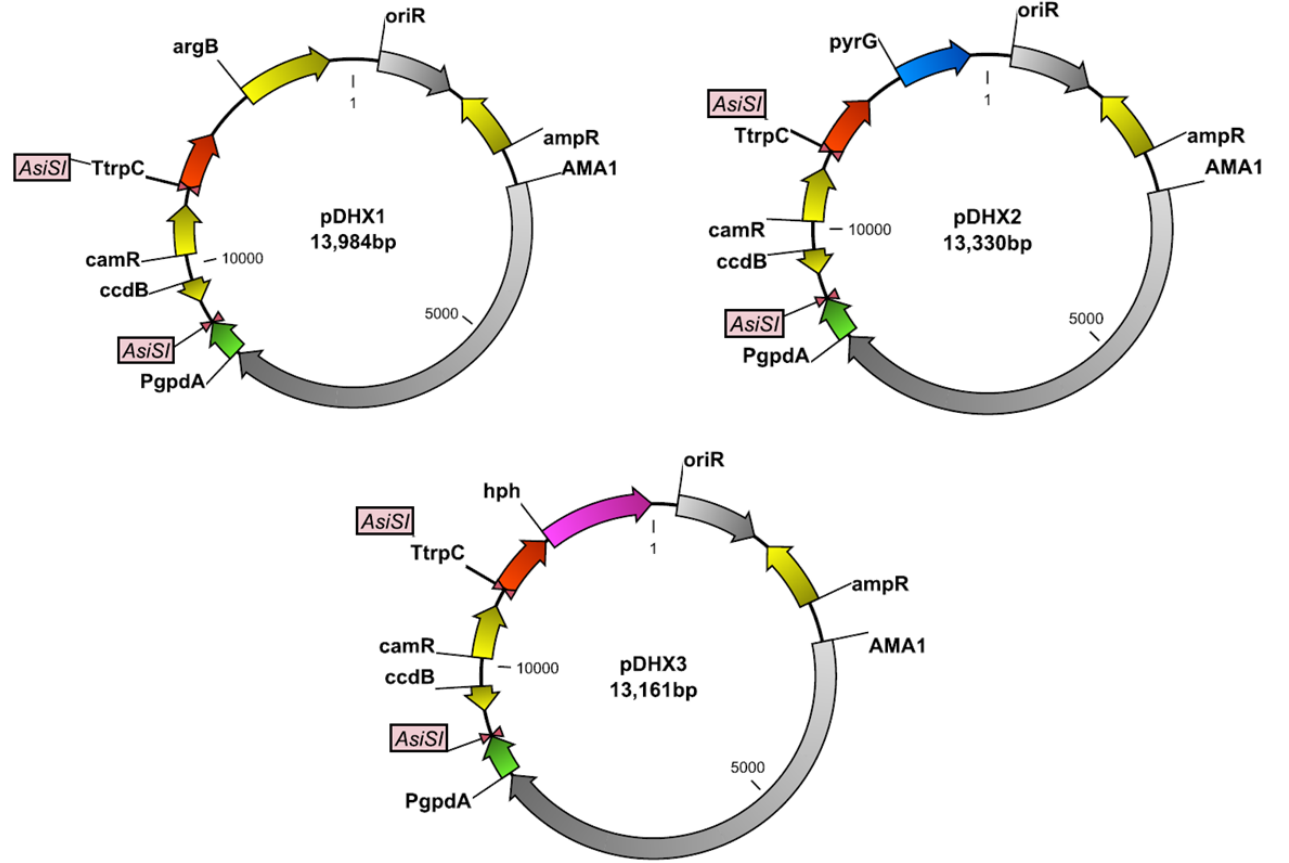


Figure 3.6: AMA1-based expression vectors developed in this study. pDHX1 contains an *argB* selectable marker, pDHX2 contains *pyrG*, and pDHX3 contains *hph*.

There are advantages and disadvantages of expressing genes from a plasmid versus integrating the expression cassette into the genome. One of the main advantages is that AMA1 plasmids have a higher transformation frequency than linear substrates [Gems et al., 1991], and thus it is easier to obtain transformants - even for strains or species that make poor competent cells e.g. protoplasts. In continuation of this, it is possible to transform with different plasmids carrying different selection markers, and obtain transformants that contain all of these plasmids. For targeted integration it is difficult to integrate more than one substrate, and at the least it requires excellent protoplasts.

On the other hand, the use of AMA1 plasmids have several drawbacks: plasmids are easily lost if selection is not maintained, and the copy number of the plasmids vary throughout the mycelium. To test the behavior

of the AMA1 plasmids, I transformed a $wA\Delta yA\Delta$ (white) strain of *A. nidulans* with two different plasmids; one carrying wA and one carrying yA (see Figure 3.7).

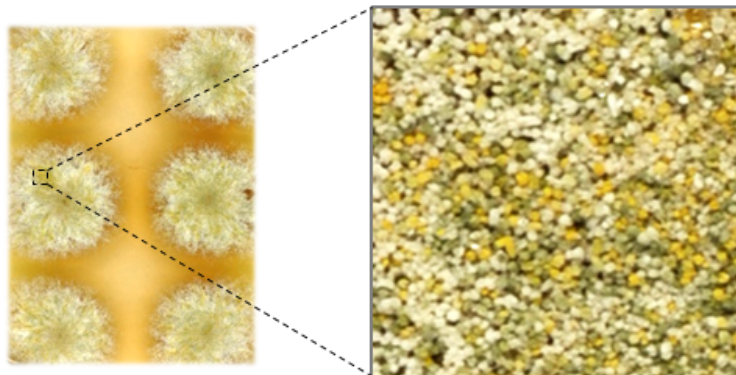


Figure 3.7: Test of AMA1 expression vectors using a color screen: $wA\Delta yA\Delta$ (white) strain transformed with pDHX1- yA (*argB*) and pDHX2- wA (*pyrG*) results in production of yellow and green conidia.

Clearly, both wA and yA were expressed as the strain was able to produce green conidia (consult Figure 1.6 for explanation). As expected, the strain produces a mixture of white, yellow, and green conidia, indicating that the plasmids are not distributed evenly throughout the nuclei. However, this tool is very suitable for “quick and dirty” screening purposes due to the ease of strain construction.

3.1.5 Development of a HTP gene expression platform for *A. nidulans*

In the following I will describe two approaches for optimizing HTP fungal strain construction; both with the aim of eliminating or drastically decrease the amount of screening. The entire HTP setup has been applied and is described in Chapter 6, why this section will provide only a brief introduction to the novel background-free cloning vectors and the background-free gene expression system for *A. nidulans*.

Development of background-free cloning vectors

One of the great challenges of HTP strain construction is to eliminate or highly reduce the amount of screening at each step. One of the bottlenecks I identified during semi-HTP PKS expression (Chapter 8) was screening for correct *E. coli* clones. The false positives background of the existing UE-compatible vectors could result from either incomplete digestion of the vector or from religation of a fully digested vector. I had observed that some of the cassettes had a higher rate of false positives than others, why I hypothesized that the false positives primarily were due to incomplete digestion by the restriction enzymes. Based on this hypothesis, I wished to develop a vector which would only be viable if completely digested.

Inspired by the background-free GatewayTM Cloning system (Life Technologies/Invitrogen), which was used in our laboratory for a while, I wondered if it was possible to transfer the *ccdB* suicide gene from the GatewayTM Cloning system into the UE-compatible vectors in a meaningful manner. Thus I inserted the

ccdB gene, containing 3' and 5' flanking sequences to reconstitute the *AsiSI*/*Nb.BtsI* cassette on either side of the suicide gene, into the cloning cassettes of CMBU1111, resulting in the vector pDH56 (see figure 3.8). The vectors were tested and proved to have extremely little background (<1 percent)(data not shown). Therefore, all of the vectors I have made in this study contains the *ccdB* suicide gene in between two identical (inverted) UE-cassettes, as shown in Figure 3.8. Digestion of pDH56 with *AsiSI* and *Nb.BtsI* releases the *ccdB::camR* cassette, leaving 8 bp single stranded overhangs in both ends. The digested pDH56 is identical to the fully digested CMBU1111 vector.

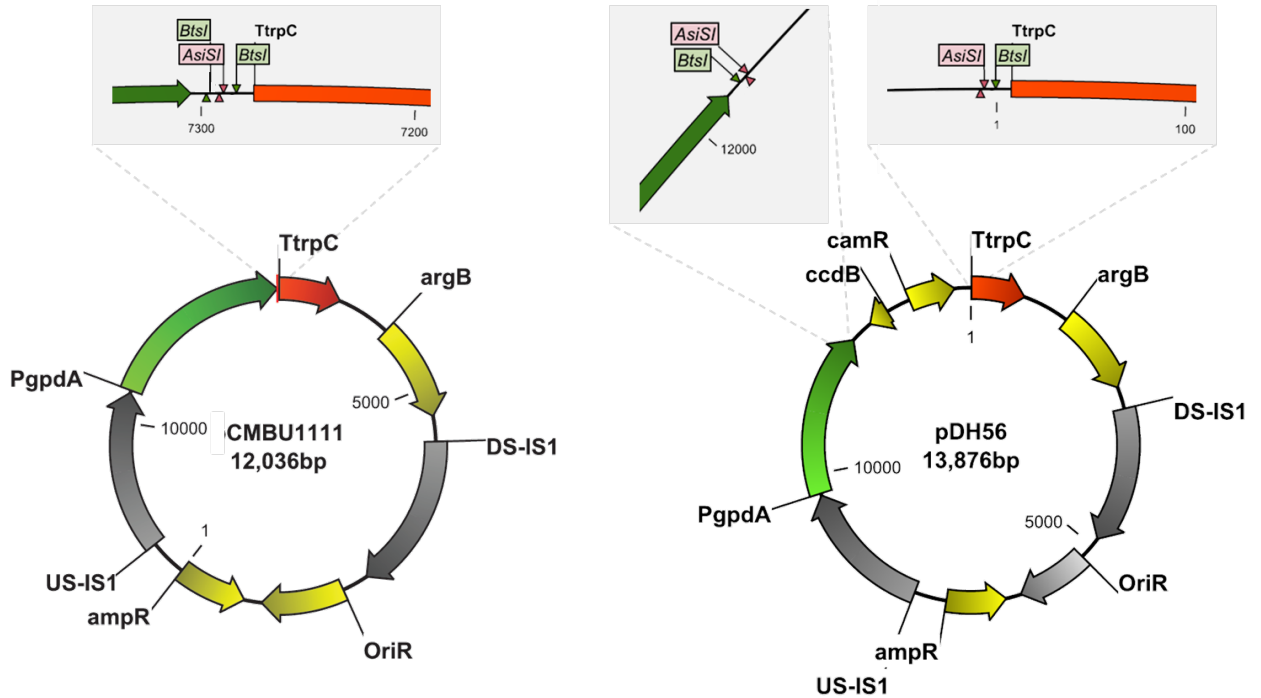


Figure 3.8: Schematic of the original expression vector CMBU1111 [Hansen et al., 2011], and the new background-free expression vector pDH56, containing the suicide gene *ccdB*.

Development of a background-free integration system

During fungal strain construction, screening for correct fungal transformants is by far the predominant bottleneck. In most cases, several transformants must be purified and re-cultivated for extraction of genomic DNA, which is used for verification of the strain by diagnostic PCR and/or southern blot. For a few strains this task is manageable, but for many strains the process is tedious and laborious. For targeted gene expression, I have developed a system that employs visual detection of false positive transformants, and even elimination of false positives by using counter-selectable markers.

The principle of the system is to prepare a primary strain that will serve as recipient of the genes that one wishes to express (Figure 3.9). The primary strain is constructed by integration of the *wA* gene, using *pyrG* as a selection marker (Figure 3.9-A). The primary strain is then used for transformation with *your favorite*

gene (YFG), which is any gene that one wishes to express (Figure 3.9-B), and this construct contains a different selection marker (here: *argB*). This construct will replace the existing cassette and loop out *wA* and *pyrG* by direct repeat recombination. Correct transformants will no longer produce the yellow compound YWA1, and are easily recognizable on the transformation plate (see Figure 2.1). Addition of 5-FOA to the transformation medium will kill all cells that express *pyrG*, thus eliminating false positive transformants. For integration sites with medium or high expression levels, addition of 5-FOA to the transformation medium is not essential, as the clear color shift from yellow to wild-type is easily distinguishable on the transformation plate.

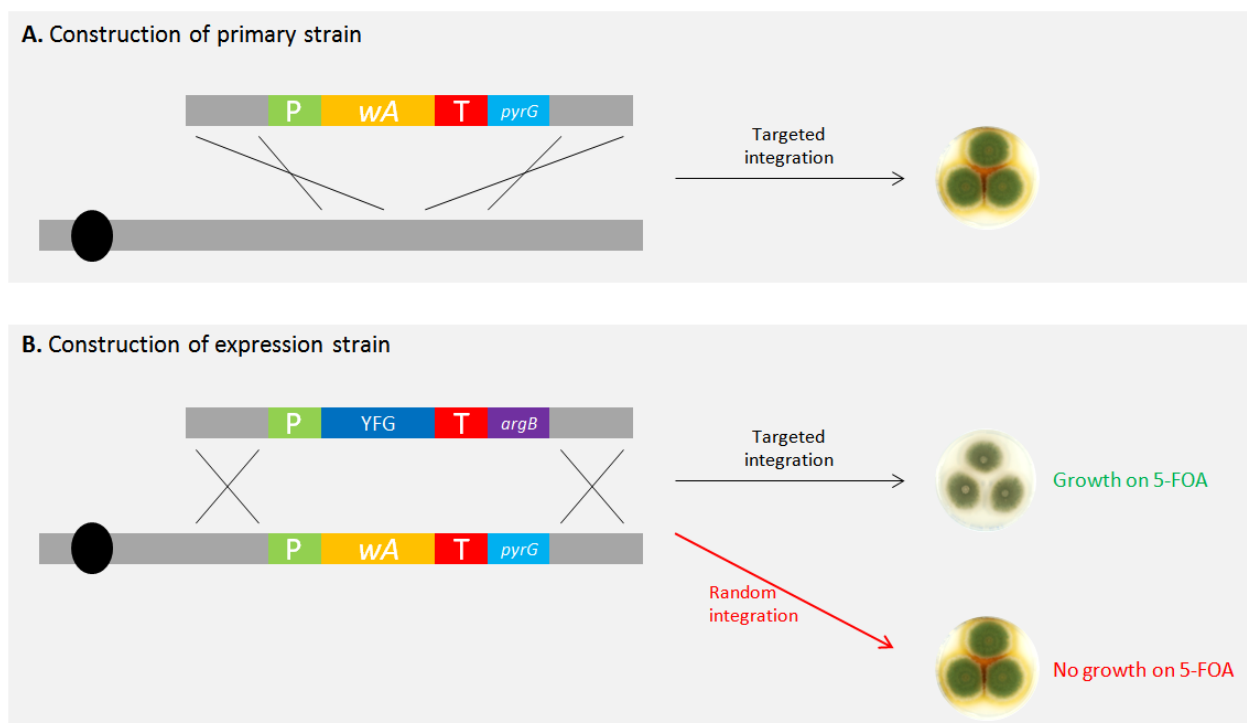


Figure 3.9: Background-free transformation system in *A. nidulans*. **A.** The primary (recipient) strain, which will be used for all future transformations, is constructed by targeted integration of the *wA* gene, which is responsible for production of the yellow pigment YWA1. The construct carries a selectable marker, in this case *pyrG*. **B.** The primary (recipient) strain is transformed with a construct carrying *your favorite gene* (YFG) and a second selectable marker, here *argB*. By targeted integration, the *wA* and *pyrG* cassette is looped out, and the pigment is no longer produced. For elimination of background, 5-FOA can be added to the transformation medium, killing all cells that still express *pyrG*.

This system is truly efficient, and I have successfully used the system for HTP site-directed mutagenesis in a commercial protease (not presented in this thesis), and for HTP over-expression of *A. nidulans* phosphatases and transcription factors (described in Chapter 6).

3.2 Concluding remarks

In this study I have sought to improve existing tools and develop new tools for genetic engineering in *Aspergilli*. I have identified six novel promoters from *A. nidulans* which, based on a qualitative X-gal assay, appear to be strongly and constitutively expressed under laboratory conditions. Two of these comprise a short (200 bp) bidirectional promoter, providing very high expression levels in either direction.

Furthermore, I have developed five new integration sites for *A. nidulans*, making a total of eight defined integration sites, including the *yA* locus, as applied by Chiang and co-workers [Chiang et al., 2013]. One of the novel integration sites, IS5, proved extremely efficient and expression of PKSs from this site provides vast amounts of secondary metabolites (e.g. 6-MSA and YWA1). Additionally, I have identified four integration sites for *A. niger*, one of which targets the *albA* gene, responsible for black pigmentation of conidia. In total, nine novel integration sites were developed in this study, allowing for ectopic or heterologous expression of several genes in the same strain.

An important result of this study was development of an AMA1-based gene expression platform. AMA1 allows for autonomous replication in several *Aspergilli*, and is independent of genetic tools for genomic integration. Furthermore, AMA1-based plasmids have a much higher transformation efficiency than linear substrates [Gems et al., 1991], why this method is also valuable for fungi that transform poorly, whether be it by protoplastation or electroporation.

Last, I have developed a system for HTP gene integration, comprising a concept for background-free cloning vectors and a background-free integration system for *A. nidulans*. The background-free cloning vectors contain a suicide gene (*ccdB*) which is excised when a fragment is correctly inserted. This system has been applied for all of the cloning vectors prepared during this study. The background-free integration system uses a recipient strain carrying a counter-selectable *pyrG* marker and a PKS gene producing the yellow compound YWA1. Correct integration will excise the *wA::pyrG* cassette, and counter-selection for *pyrG* with 5-FOA will allow growth of only correct transformants. It is possible to use any markers if the PKS is highly expressed, as the color-shift makes it easy to verify correct transformants. Thus, this system should also be fully applicable onto other filamentous fungi.

3.3 Experimental

3.3.1 Fungal strains

The *A. nidulans* strain NID1 (IBT 29539) (*argB2 pyrG89 veA1 nkuAΔ*) [Nielsen et al., 2008] was used as a reference and for insertion of all genes. For *A. niger*, KB1001 (*pyrGΔ kusA::AFpyrG*) [Chiang et al., 2011] was used as a reference and for genetic engineering.

Table 3.4: Fungal strains used and constructed in this study. *hph** contains a silent point mutation eliminating an AsiSI restriction site. *ori^R** contains a point mutation eliminating a Nb.BtsI nicking site. FGSC = Fungal Genetics Stock Center, CMB SC = CMB *Aspergillus* Strain Collection.

Strain	Genotype	Source
<i>A. nidulans</i> strains		
A4 (FGSC)	-	FGSC
NID1	<i>argB2 pyrG89 nkuAΔ</i>	[Nielsen et al., 2008]
Prom1	<i>argB2 pyrG89 nkuAΔ</i> IS1::P _{AN0262} :: <i>lacZ</i> :: <i>TtrpC</i> :: <i>argB</i>	This study
Prom2	<i>argB2 pyrG89 nkuAΔ</i> IS1::P _{AN1122} :: <i>lacZ</i> :: <i>TtrpC</i> :: <i>argB</i>	This study
Prom3	<i>argB2 pyrG89 nkuAΔ</i> IS1::P _{AN4794} :: <i>lacZ</i> :: <i>TtrpC</i> :: <i>argB</i>	This study
Prom4	<i>argB2 pyrG89 nkuAΔ</i> IS1::P _{AN4802} :: <i>lacZ</i> :: <i>TtrpC</i> :: <i>argB</i>	This study
Prom5	<i>argB2 pyrG89 nkuAΔ</i> IS1::P _{AN4803} :: <i>lacZ</i> :: <i>TtrpC</i> :: <i>argB</i>	This study
Prom6	<i>argB2 pyrG89 nkuAΔ</i> IS1::P _{AN5997} :: <i>lacZ</i> :: <i>TtrpC</i> :: <i>argB</i>	This study
Prom7	<i>argB2 pyrG89 nkuAΔ</i> IS1::P _{AN6082} :: <i>lacZ</i> :: <i>TtrpC</i> :: <i>argB</i>	This study
Prom8	<i>argB2 pyrG89 nkuAΔ</i> IS1::P _{AN7354} :: <i>lacZ</i> :: <i>TtrpC</i> :: <i>argB</i>	This study
Prom9	<i>argB2 pyrG89 nkuAΔ</i> IS1::P _{AN8704} :: <i>lacZ</i> :: <i>TtrpC</i> :: <i>argB</i>	This study
Prom10	<i>argB2 pyrG89 nkuAΔ</i> IS1::P _{AN8856} :: <i>lacZ</i> :: <i>TtrpC</i> :: <i>argB</i>	This study
Prom11	<i>argB2 pyrG89 nkuAΔ</i> IS1::P _{AN10130} :: <i>lacZ</i> :: <i>TtrpC</i> :: <i>argB</i>	This study
NID545	<i>argB2 pyrG89 nkuAΔ</i> IS1::P _{gpdA_{0.5kb}} :: <i>lacZ</i> :: <i>TtrpC</i> :: <i>pyrG</i>	CMB SC
IS3- <i>lacZ</i>	<i>argB2 pyrG89 nkuAΔ</i> IS3::P _{gpdA_{0.5kb}} :: <i>lacZ</i> :: <i>TtrpC</i> :: <i>pyrG</i>	This study
IS4- <i>lacZ</i>	<i>argB2 pyrG89 nkuAΔ</i> IS4::P _{gpdA_{0.5kb}} :: <i>lacZ</i> :: <i>TtrpC</i> :: <i>pyrG</i>	This study
IS5- <i>lacZ</i>	<i>argB2 pyrG89 nkuAΔ</i> IS5::P _{gpdA_{0.5kb}} :: <i>lacZ</i> :: <i>TtrpC</i> :: <i>pyrG</i>	This study
IS6- <i>lacZ</i>	<i>argB2 pyrG89 nkuAΔ</i> IS6::P _{gpdA_{0.5kb}} :: <i>lacZ</i> :: <i>TtrpC</i> :: <i>pyrG</i>	This study
IS7- <i>lacZ</i>	<i>argB2 pyrG89 nkuAΔ</i> IS7::P _{gpdA_{0.5kb}} :: <i>lacZ</i> :: <i>TtrpC</i> :: <i>pyrG</i>	This study
IS1- <i>wA</i>	<i>argB2 pyrG89 nkuAΔ</i> IS1::P _{gpdA_{0.5kb}} :: <i>wA</i> :: <i>TtrpC</i> :: <i>pyrG</i>	This study
IS3- <i>wA</i>	<i>argB2 pyrG89 nkuAΔ</i> IS3::P _{gpdA_{0.5kb}} :: <i>wA</i> :: <i>TtrpC</i> :: <i>pyrG</i>	This study
IS4- <i>wA</i>	<i>argB2 pyrG89 nkuAΔ</i> IS4::P _{gpdA_{0.5kb}} :: <i>wA</i> :: <i>TtrpC</i> :: <i>pyrG</i>	This study
IS5- <i>wA</i>	<i>argB2 pyrG89 nkuAΔ</i> IS5::P _{gpdA_{0.5kb}} :: <i>wA</i> :: <i>TtrpC</i> :: <i>pyrG</i>	This study

Continued on Next Page

Table 3.4 – Continued

Strain	Genotype	Source
IS6- <i>wA</i>	<i>argB2 pyrG89 nkuAΔ</i> IS6::P <i>gpdA</i> _{0.5kb} :: <i>wA</i> ::T <i>trpC</i> :: <i>pyrG</i>	This study
IS7- <i>wA</i>	<i>argB2 pyrG89 nkuAΔ</i> IS7::P <i>gpdA</i> _{0.5kb} :: <i>wA</i> ::T <i>trpC</i> :: <i>pyrG</i>	This study
<i>wAΔ yAΔ</i>	<i>argB2 pyrG89 nkuAΔ wAΔ yAΔ</i>	This study
2xAMA1	<i>argB2 pyrG89 nkuAΔ</i> pDHX1- <i>yA</i> pDHX2- <i>wA</i>	This study
<i>A. niger</i> strains		
KB1001	<i>pyrGΔ kusA::AFpyrG</i>	PNNL
IS-NIG2- <i>rfp</i>	<i>pyrGΔ kusA::AFpyrG</i> IS-NIG2::P <i>gpdA</i> _{0.5kb} :: <i>rfp</i> ::T <i>trpC</i> :: <i>hph</i> *	This study
IS-NIG3- <i>wA</i>	<i>pyrGΔ kusA::AFpyrG</i> IS-NIG3::P <i>gpdA</i> _{0.5kb} :: <i>wA</i> ::T <i>trpC</i> :: <i>hph</i> *	This study
IS-NIG4- <i>wA</i>	<i>pyrGΔ kusA::AFpyrG</i> IS-NIG4::P <i>gpdA</i> _{0.5kb} :: <i>wA</i> ::T <i>trpC</i> :: <i>hph</i> *	This study

3.3.2 Media

Minimal medium for *A. nidulans* was made as described in [Cove, 1966] but with 1% glucose, 10 mM NaNO₃ and 2% agar. Minimal medium for *A. niger* was prepared as described in [Chiang et al., 2011]. YES, MEA and CYA media were prepared as described by Frisvad and Samson [Samson et al.]. When necessary, media were supplemented with 0.122 mM 5-bromo-4-chloro-3-indoyl- β -D-galactopyranoside (X-gal), 4 mM L-arginine, 10 mM uridine, 10 mM uracil and/or 100 μ g/ml hygromycin B (Invivogen, San Diego, CA, USA) when necessary.

3.3.3 PCR, cloning and transformation

All primers are listed in Table 3.3.3. All PCR reactions were carried out using the *PfuX7* polymerase [Nørholm, 2010], and either HF or Phire PCR buffer (Finnzymes/Thermo Scientific). The PCR program was: 98°C for 1 min, 98°C for 15 s, 60°C for 30 s, 72°C for 30s/kb, and 72°C for 5 min. Normally, I performed 35 cycles of amplification. Diagnostic (conidial) PCR was performed as a regular PCR reaction, except initial denaturation step was 20 min.

Fragments were assembled via uracil-excision fusion [Geu-Flores et al., 2007] into a compatible vector. Correct clones were verified by restriction digestion. Protoplasting and gene-targeting procedures were performed as described previously [Johnstone et al., 1985, Nielsen et al., 2006]. pDH56 was prepared as described in Chapter 10.

Construction of strains with new promoters

Promoter regions were amplified using primers Prom-X-FW and Prom-X-RV, where X is the locus of the gene downstream of the promoter (primers 1-22 in Table 3.3.3). The promoter fragments were inserted into plasmid p68 as described in [Hansen et al., 2011]. Plasmids were verified by restriction analysis, linearized with NotI and transformed into NID1. Correct transformants were verified by diagnostic PCR as described in [Hansen et al., 2011].

Construction of plasmids with new integration sites

Plasmids for integration into new sites were constructed by uracil excision (UE) fusion of five fragments: (1) one fragment comprising *E. coli* origin of replication *ori^R* and ampicillin resistance gene *amp^R*, (2) the upstream (US) targeting sequence, (3) one fragment comprising *PgpdA_{0.5kb}*, UE cassette containing *ccdB* and *cm^R*, and *TtrpC*, (4) marker, and (5) downstream (DS) targeting sequence.

Fragments 1 and 3 were amplified from pDH57 (primers 52+53 and primers 54+55, respectively) 10. Fragment 4 (marker) was amplified from pDH57 (*argB*, primers 57+58), CMBU2111 (*pyrG*, primers 59+60 or 61+62) [Hansen et al., 2011], or pCB1003 (*hph*, primers 98+99) [McCluskey et al., 2010]. US and DS fragments were amplified from *A. nidulans* or *A. niger* genomic DNA, using primers 25-28 (IS7), primers 29-32 (IS4), primers 33-36 (IS3), primers 37-40 (IS5), primers 41-44 (IS6), primers 67-70 (IS-NIG1), primers 73-76 (IS-NIG2), primers 77-80 (IS-NIG3), or primers 83-86 (IS-NIG4).

Construction of strains for testing new integration sites

Integration sites were tested by expression of *lacZ* and *wA*, respectively, or *rfp*. *lacZ* was PCR amplified from p68 using primers 100 and 101, *wA* was PCR amplified from *A. nidulans* IBT29539 gDNA using primers 102 and 103, and *rfp* was amplified from CMBU2111-*rfp* (kindly provided by Diana C. Anyaogu) using primers 92+93. The PCR fragments were inserted into digested plasmids (described above), verified, linearized by digestion with *SwaI* (for *pyrG* containing plasmids), and transformed into *A. nidulans* IBT29539 or *A. niger* KB1001. Correct transformants were verified by diagnostic PCR.

Construction of strain for testing expression from AMA1 plasmids

The 2xAMA1 strain was constructed for testing the expression platform in the novel AMA1-based plasmids. *wA* and *yA* were PCR amplified from *A. nidulans* IBT29539 gDNA (primers 102+103 and 104+105) and inserted into pDHX2, and pDHX1, respectively. Clones were verified and transformed directly into the *wAΔ yAΔ* strain as described above.

Table 3.5: Primers used in this study. All primers are shown in 5'-3' direction

Primer number	Primer name	Sequence
1	Prom-AN5997-FW	AGAGCGAUAGAAGGAGGAGTCCGATG
2	Prom-AN5997-RV	TCTGCGAUGGATGGCGTGTGCTGTT
3	Prom-AN8856-FW	AGAGCGAUATCGAGGTGTTTGTGCTGA
4	Prom-AN8856-RV	TCTGCGAUTGTGTCGAGTGTGCGAAAGTCG
5	Prom-AN4794-FW	AGAGCGAUCAGGCCTTTTGTGTTGGAC
6	Prom-AN4794-RV	TCTGCGAUTGTGTCGCGAAGAAAAGAAGC
7	Prom-AN7354-FW	AGAGCGAUGAGCTTTGTGCTGGAGTT
8	Prom-AN7354-RV	TCTGCGAUGTCGAAATCGGTCTCGAAATT
9	Prom-AN6082-FW	AGAGCGAUGATGGCGCTCATATCGGT
10	Prom-AN6082-RV	TCTGCGAUATGAGATGCAAAAAGACGGT
11	Prom-AN1122-FW	AGAGCGAUATCTCACTCCACTAGAATTCCT
12	Prom-AN1122-RV	TCTGCGAUAATGAGTCGAGAATGGCG
13	Prom-AN0262-FW	AGAGCGAUTCCAAGCTTAACCCGTCC
14	Prom-AN0262-RV	TCTGCGAUAGAGGAGGGAAGATTCGAG
15	Prom-AN8704-FW	AGAGCGAUGGGGTAGGTGCTAGATGTC
16	Prom-AN8704-RV	TCTGCGAUGACCGGTTGGGACGAGATG
17	Prom-AN4802-FW	AGAGCGAUCGAATAACTCAGCTGAACGACTT
18	Prom-AN4802-RV	TCTGCGAUGAAAACGAAGAAGTCCCTCGAT
19	Prom-AN4803-FW	AGAGCGAUGAAAACGAAGAAGTCCCTCGAT
20	Prom-AN4803-RV	TCTGCGAUCGAATAACTCAGCTGAACGACTT
21	Prom-AN10130-FW	AGAGCGAUTGCATCTTTCGAGTCCCTATTT
22	Prom-AN10130-RV	TCTGCGAUGTTGTTGAGAGCTGCGAG
23	ccdB-camR-fw	AGAGCGAUCGCAGAAGCCTACTCGCTATTGTCCTCA
24	ccdB-camR-rv	TCTGCGAUCGCTCTTGCGCCGAATAAATACCTGT
25	AN8282-US-Fw	AAGCAGCUATCGGTCATGTGATGCTGAGG
26	AN8282-US-Rv	ATATTGGGCTGCTTGGGAATATTATGATGATTCAGC
27	AN8282-DS-Fw	ACTAGGTAAUCTTCGGACTTATCAACCCAGTG
28	AN8282-DS-Rv	AGTGGGGAUAGGCAGGCCAATCGACTAA
29	AN4252-US-Fw	AAGCAGCUCTCAGAGCATCTTTCTCTACCCAT
30	AN4252-US-Rv	ATATTGGGCTGGACCGCATCTGATTGAACG
31	AN4252-DS-Fw	ACTAGGTAAUCTAGTTGTCAGTAGTACGGCAAGC
32	AN4252-DS-Rv	AGTGGGGAUTCCAATGGTAGGGACATTGTTG

Continued on Next Page

Table 3.5 – Continued

Primer number	Primer name	Sequence
33	AN4770-US-Fw	AAGCAGCUCAAGAGGTACCCGCTACAGCA
34	AN4770-US-Rv	ATATTGGGCUAGCCTTCATTGACCAGCTTCG
35	AN4770-DS-Fw	ACTAGGTAAUTGAATGTTGGAATAGGCTGTGC
36	AN4770-DS-Rv	AGTGGGGAUCTGCATACCCGTCGTAGTCCT
37	AN7553-US-Fw	AAGCAGCUGTTTGCCATGCTGCAACTGT
38	AN7553-US-Rv	ATATTGGGCUTCCTGAGTGTGAGTCTGACTTC
39	AN7553-DS-Fw	ACTAGGTAAUGCATGGCAATCAAGTCCCT
40	AN7553-DS-Rv	AGTGGGGAUAGGGAACGGTTGTCTGATGG
41	AN3180-US-Fw	AAGCAGCUTAAGGGAGACCGTCAGAGTTTG
42	AN3180-US-Rv	ATATTGGGCUACGTCTCCTGCAGAGTAGCTGT
43	AN3180-DS-Fw	ACTAGGTAAUCAAGCTAGACTGTCAAGGGACATT
44	AN3180-DS-Rv	AGTGGGGAUTGCTGAGGATGCTGAGATGAAG
45	IS4252-chk-RV	GAGAGACAACCAGCTTTCACCA
46	IS7553-chk-RV	GGTGAATACTGGAATGCGCTAAT
47	IS4770-chk-RV	TCGTCAATTGTAAGAATAGCAAGGT
48	IS3180-chk-RV	AGAACCGGGTTCTGTGGATG
49	IS8282-chk-RV	TCTGAAACTCACCGTCAGTGCTTA
50	ampR-PM-FW	AGCGCTACAUAATTCTCTTACTGTCATGCCATCC
51	ampR-PM-RV	ATGTAGCGCUGCCATAACCATGAGTGATAACACTG
52	ori-coli-FW	ATCCCCACUACCGCATTAAGACCTCAGCG
53	ampR-RV	AGCTGCTUCGTGCTGATTAAACCCCTCAGCG
54	p71-prom-ter-short-usF	AGCCCAATAUTAAGCTCCCTAATTGGCCC
55	p71-prom-ter-dsR	ATTACCTAGUGGGCGCTTACACAGTACA
56	p71-prom-ter-IS-dsR	AGCTCACUGGGCGCTTACACAGTACA
57	argB-FW	ACTAGGTAAUATCGCGTGCAATCCGCGGT
58	argB-RV-forIS	ATTACCTAGUGCCTCAGCGCCACCTACA
59	pyrG-FW-C1	ACTAGGTAAUATGACATGATTACGAATTCGAGCT
60	pyrG-RV-forIS	ATTACCTAGUGCCTCAATTGTGCTAGCTGC
61	pyrG-FW-DR-forIS	AGTGAGCUTGGATAACCGTATTACCGCC
62	pyrG-RV-DR-forIS	ATTACCTAGUTGCCAAGCTTAACGCGTAC
63	amdS-fw	ACTAGGTAAUTGGATAGGCAGATTACTCAGCCT
64	amdS-rv	AGTGGGGAUGGACAGTCTGCACGTTACTGCT
65	BleR-fw	ACTAGGTAAUGCTGATTCTGGAGTGACCCAG

Continued on Next Page

Table 3.5 – Continued

Primer number	Primer name	Sequence
66	BleR-rv	AGTGGGGAUCGTTGTAAAACGACGGCC
67	ASN-IS1-US-Fw	AAGCAGCUACGCTTGGATAAGAAGTGAGAAGT
68	ASN-IS1-US-Rv	ATATTGGGCUTCTACACCGATATGGAGCTCTTT
69	ASN-IS1-DS-Fw	ACTAGGTAAUAAGTAAGCGTGC GTGAAACAG
70	ASN-IS1-DS-Rv	AGTGGGGAUTAATTATAACTGGTCAGCCACTGAC
71	Chk-ASNIS1-US-Fw	GGAGTTTCGGTGGAGTCTACGGAG
72	Chk-ASNIS1-DS-Rv	TGGGTTACATACCTCTTCTGGGT
73	AlbA-US-FW	AAGCAGCUATCCGTCCGTTCAGGCTGCT
74	AlbA-US-RV	ATATTGGGCUCACACGAGATGGACCCTCCAT
75	AlbA-DS-FW	ACTAGGTAAUTCCCGCTAGGACTGACAACA
76	AlbA-DS-RV	AGTGGGGAUCTCCCTCGTATTCGCCTCCT
77	ISU3-US-Fw	AAGCAGCUAGTTAGGGACTTAGCCTCGGAC
78	ISU3-US-Rv	ATATTGGGCUCTCTTCGAGACTGGCTAGATAGATAG
79	ISU3-DS-Fw	ACTAGGTAAUCACTAGACAGGTTGAAGGTACTTTGT
80	ISU3-DS-Rv	AGTGGGGAUTC GCATGTACAGTGGTCTCTTATG
81	Chk-ISU3-Fw	TTAGTTCATGGATGCTCGCCT
82	Chk-ISU3-Rv	TCAGACAGCATGATAACAGCCA
83	ISU4-US-Fw	AAGCAGCUTACTTGGTGGCTGTAAGAGTATGACTA
84	ISU4-US-Rv	ATATTGGGCUTGTGGCACTGGTGTCTCGTC
85	ISU4-DS-Fw	ACTAGGTAAUGTCTGGTTCCTGTGAACTAACTAGAC
86	ISU4-DS-Rv	AGTGGGGAUTCATGTTTCAGGGATCGGCA
87	ISU4-DS-Rv2	AGTGGGGAUAGTATTCTGATGTGTTTCGAGACCA
88	Chk-ISU4-Fw	CATGTTGGAAAGTCTGATGATAATAATC
89	Chk-ISU4-Rv	GTA CTCTGAGGCATGGACCAAT
90	hph-PM-Fw	ATGCGATUGCTGCGGCCGATCTTAGC
91	hph-PM-Rv	AATCGCAUCCATGGCCTCCGCGAC
92	mRFP-ORF-Fw	AGAGCGAUATGGCCTCCTCCGAGGAC
93	mRFP-ORF-Rv	TCTGCGAUTTAGGCGCCGGTGGAGT
94	AMA1-FW	AAGCAGCUGACGGCCAGTGCCAAGCT
95	AMA1-RV	ATATTGGGCUGGAAACAGCTATGACCATGAGATCT
96	AMA-3'-Fw	ACCCCAAUGGAAACGGTGAGAGTCCAGTG
97	AMA-5'-Rv	ATTGGGGUACTAACATAGCCATCAAATGCC
98	hph-1003-pX-Fw	ACTAGGTAAUGCTAGTGGAGGTCAACACATCA

Continued on Next Page

Table 3.5 – Continued

Primer number	Primer name	Sequence
99	hph-1003-pX-Rv	AGTGGGGAUCGGTCGGCATCTACTCTATT
100	lacZ-Fw	AGAGCGAUATGACCATGATTACGGATTC
101	lacZ-Rv	TCTGCGAUTTATTTTTTGACACCAGACCA
102	AN8209-Fw	AGAGCGAUATGGAGGACCCATACCGTGT
103	AN8209-Rv	TCTGCGAUTATTAGAACAGAGGATTATTATTGTT
104	AN6635-Fw	AGAGCGAUATGTACCTCTCCACGGTCCT
105	AN6635-Rv	TCTGCGAUCTAAGAATCCCAAACATCAAC

Table 3.6: Plasmids used in this study. All plasmids aer *ori^R amp^R*. *hph** contains a silent point mutation eliminating an AsiSI restriction site. *ori^R** contains a point mutation eliminating a Nb.BtsI nicking site. UEC = AsiSI/Nb.BtsI uracil excision cassette

Plasmid	Genotype	Source
p68	TS1::UEC::lacZ::TtrpC::argB::TS2	
CMBU1111	TS1::P _{gpdA} ::UEC::TtrpC::argB::TS2	[Hansen et al., 2011]
CMBU1211	TS1::P _{alcA} ::UEC::TtrpC::argB::TS2	[Hansen et al., 2011]
CMBU2111	TS1::P _{gpdA} ::UEC::TtrpC::pyrG::TS2	[Hansen et al., 2011]
pProm1	TS1::P _{AN0262} ::lacZ::TtrpC::argB::TS2	This study
pProm2	TS1::P _{AN1122} ::lacZ::TtrpC::argB::TS2	This study
pProm3	TS1::P _{AN4794} ::lacZ::TtrpC::argB::TS2	This study
pProm4	TS1::P _{AN4802} ::lacZ::TtrpC::argB::TS2	This study
pProm5	TS1::P _{AN4803} ::lacZ::TtrpC::argB::TS2	This study
pProm6	TS1::P _{AN5997} ::lacZ::TtrpC::argB::TS2	This study
pProm7	TS1::P _{AN6082} ::lacZ::TtrpC::argB::TS2	This study
pProm8	TS1::P _{AN7354} ::lacZ::TtrpC::argB::TS2	This study
pProm9	TS1::P _{AN8704} ::lacZ::TtrpC::argB::TS2	This study
pProm10	TS1::P _{AN8856} ::lacZ::TtrpC::argB::TS2	This study
pProm11	TS1::P _{AN10130} ::lacZ::TtrpC::argB::TS2	This study
pDH56	TS1::P _{gpdA} ::UEC::ccdB::cm ^R ::UEC::lacZ::TtrpC::argB::TS2	This study
pDH57	TS1::P _{gpdA} ::UEC::ccdB::cm ^R ::UEC::lacZ::TtrpC::argB::TS2	This study
pDH58	TS1::P _{alcA} ::UEC::ccdB::cm ^R ::UEC::lacZ::TtrpC::argB::TS2	This study
pDH59	TS1::P _{alcA} ::UEC::ccdB::cm ^R ::UEC::lacZ::TtrpC::argB::TS2	This study

Continued on Next Page

Table 3.6 – Continued

Plasmid	Genotype	Source
pDHU1	TS-NIG1A::P <i>gpdA</i> _{0.5kb} ::UEC::ccdB::cm ^R ::UEC::TtrpC::hph*::TS-NIG1B	This study
pDHU2	TS-NIG2A::P <i>gpdA</i> _{0.5kb} ::UEC::ccdB::cm ^R ::UEC::TtrpC::hph*::TS-NIG2B	This study
pDHU2002	TS-NIG2A::P <i>gpdA</i> _{0.5kb} ::rfp::TtrpC::hph*::TS-NIG2B	This study
pDHU3	TS-NIG3A::P <i>gpdA</i> _{0.5kb} ::UEC::ccdB::cm ^R ::UEC::TtrpC::hph*::TS-NIG3B	This study
pDHU3000	TS-NIG3A::P <i>gpdA</i> _{0.5kb} ::lacZ::TtrpC::hph*::TS-NIG3B	This study
pDHU4	TS-NIG4A::P <i>gpdA</i> _{0.5kb} ::UEC::ccdB::cm ^R ::UEC::TtrpC::hph*::TS-NIG4B	This study
pDHU4000	TS-NIG4A::P <i>gpdA</i> _{0.5kb} ::lacZ::TtrpC::hph*::TS-NIG4B	This study
pDHA3	TS3A::P <i>gpdA</i> _{0.5kb} ::UEC::ccdB::cm ^R ::UEC::TtrpC::argB::TS3B	This study
pDHP3	TS3A::P <i>gpdA</i> _{0.5kb} ::UEC::ccdB::cm ^R ::UEC::TtrpC::pyrG::TS3B	This study
pDHP3000	TS3A::P <i>gpdA</i> _{0.5kb} ::lacZ::TtrpC::pyrG::TS3B	This study
pDHP3001	TS3A::P <i>gpdA</i> _{0.5kb} ::wA::TtrpC::pyrG::TS3B	This study
pDHD3	TS3A::P <i>gpdA</i> _{0.5kb} ::UEC::ccdB::cm ^R ::UEC::TtrpC::DR-pyrG-DR::TS3B	This study
pDHA4	TS4A::P <i>gpdA</i> _{0.5kb} ::UEC::ccdB::cm ^R ::UEC::TtrpC::argB::TS4B	This study
pDHP4	TS4A::P <i>gpdA</i> _{0.5kb} ::UEC::ccdB::cm ^R ::UEC::TtrpC::pyrG::TS4B	This study
pDHP4000	TS4A::P <i>gpdA</i> _{0.5kb} ::lacZ::TtrpC::pyrG::TS4B	This study
pDHP4001	TS4A::P <i>gpdA</i> _{0.5kb} ::wA::TtrpC::pyrG::TS4B	This study
pDHD4	TS4A::P <i>gpdA</i> _{0.5kb} ::UEC::ccdB::cm ^R ::UEC::TtrpC::DR-pyrG-DR::TS4B	This study
pDHA5	TS5A::P <i>gpdA</i> _{0.5kb} ::UEC::ccdB::cm ^R ::UEC::TtrpC::argB::TS5B	This study
pDHP5	TS5A::P <i>gpdA</i> _{0.5kb} ::UEC::ccdB::cm ^R ::UEC::TtrpC::pyrG::TS5B	This study
pDHP5000	TS5A::P <i>gpdA</i> _{0.5kb} ::lacZ::TtrpC::pyrG::TS5B	This study
pDHP5001	TS5A::P <i>gpdA</i> _{0.5kb} ::wA::TtrpC::pyrG::TS5B	This study
pDHD5	TS5A::P <i>gpdA</i> _{0.5kb} ::UEC::ccdB::cm ^R ::UEC::TtrpC::DR-pyrG-DR::TS5B	This study
pDHA6	TS6A::P <i>gpdA</i> _{0.5kb} ::UEC::ccdB::cm ^R ::UEC::TtrpC::argB::TS6B	This study
pDHP6	TS6A::P <i>gpdA</i> _{0.5kb} ::UEC::ccdB::cm ^R ::UEC::TtrpC::pyrG::TS6B	This study
pDHP6000	TS6A::P <i>gpdA</i> _{0.5kb} ::lacZ::TtrpC::pyrG::TS6B	This study
pDHP6001	TS6A::P <i>gpdA</i> _{0.5kb} ::wA::TtrpC::pyrG::TS6B	This study
pDHD6	TS6A::P <i>gpdA</i> _{0.5kb} ::UEC::ccdB::cm ^R ::UEC::TtrpC::DR-pyrG-DR::TS6B	This study
pDHA7	TS7A::P <i>gpdA</i> _{0.5kb} ::UEC::ccdB::cm ^R ::UEC::TtrpC::argB::TS7B	This study
pDHP7	TS7A::P <i>gpdA</i> _{0.5kb} ::UEC::ccdB::cm ^R ::UEC::TtrpC::pyrG::TS7B	This study
pDHP7000	TS7A::P <i>gpdA</i> _{0.5kb} ::lacZ::TtrpC::pyrG::TS7B	This study
pDHP7001	TS7A::P <i>gpdA</i> _{0.5kb} ::wA::TtrpC::pyrG::TS7B	This study
pDHD7	TS7A::P <i>gpdA</i> _{0.5kb} ::UEC::ccdB::cm ^R ::UEC::TtrpC::DR-pyrG-DR::TS7B	This study

Chapter 4

A genome-wide polyketide synthase deletion library uncovers novel genetic links to polyketides and meroterpenoids in *Aspergillus nidulans*

This chapter contains the paper "A genome-wide polyketide synthase deletion library uncovers novel genetic links to polyketides and meroterpenoids in *Aspergillus nidulans*", published in *FEMS Microbiology Letters*, of which I am a co-author.

In this study we have individually deleted all putative polyketide synthases in *A. nidulans* and analyzed the metabolite profiles of the resulting strains. This resulted in identification of the polyketide synthase responsible for producing 3,5-dimethylorsellinic acid (DMOA) which proved to be the precursor to meroterpenoids austinol and dehydroaustinol. After publication of this paper, the complete biosynthesis towards these meroterpenoids was reported in detail by Itoh et al. [2012], Lo et al. [2012], and Matsuda et al. [2013].

My contribution to this study comprises over-expression of *ausA*, site-directed mutagenesis in *ausA*, preparation of large metabolite extract from IS1-*ausA*, and subsequent isolation and purification of 3,5-dimethylorsellinic acid. Supporting information is included in Appendix A.

A genome-wide polyketide synthase deletion library uncovers novel genetic links to polyketides and meroterpenoids in *Aspergillus nidulans*

Michael L. Nielsen, Jakob B. Nielsen, Christian Rank, Marie L. Klejnstrup, Dorte K. Holm, Katrine H. Brogaard, Bjarne G. Hansen, Jens C. Frisvad, Thomas O. Larsen & Uffe H. Mortensen

Department of Systems Biology, Center for Microbial Biotechnology, Technical University of Denmark, Lyngby, Denmark

Correspondence: Uffe H. Mortensen, Department of Systems Biology, Center for Microbial Biotechnology, Technical University of Denmark, Building 223, 2800 Kgs. Lyngby, Denmark. Tel.: +45 4525 2701; fax: +45 4588 4148; e-mail: um@bio.dtu.dk

Received 14 April 2011; revised 26 May 2011; accepted 27 May 2011.

DOI:10.1111/j.1574-6968.2011.02327.x

Editor: Richard Staples

Keywords

filamentous fungi; secondary metabolism; arugosin; austinol; violaceol; shamixanthone

Abstract

Fungi possess an advanced secondary metabolism that is regulated and coordinated in a complex manner depending on environmental challenges. To understand this complexity, a holistic approach is necessary. We initiated such an analysis in the important model fungus *Aspergillus nidulans* by systematically deleting all 32 individual genes encoding polyketide synthases. Wild-type and all mutant strains were challenged on different complex media to provoke induction of the secondary metabolism. Screening of the mutant library revealed direct genetic links to two austinol meroterpenoids and expanded the current understanding of the biosynthetic pathways leading to arugosins and violaceols. We expect that the library will be an important resource towards a systemic understanding of polyketide production in *A. nidulans*.

Introduction

It is well known that filamentous fungi produce a large number of bioactive secondary metabolites (Bérdy, 2005; Cox, 2007; Newman & Cragg, 2007). Polyketide (PK) compounds constitute a major part of these metabolites and have long been recognized as a valuable source of diverse natural compounds of medical importance, for example lovastatin (cholesterol lowering) (Lai *et al.*, 2005), griseofulvin (antibiotic) (Chooi *et al.*, 2010) and mycophenolic acid (immunosuppressant) (Bentley, 2000). However, polyketides also include many toxic compounds that pose a serious threat to human health, for example patulin, ochratoxins, fumonisins and aflatoxin (Frisvad *et al.*, 2004; Månsson *et al.*, 2010). Polyketides are biosynthesized by large multidomain polyketide synthases (PKSs), which besides acyl transferase, β -ketoacyl synthase and acyl carrier domains may also contain keto reductase, dehydratase, cyclization and methyl-transferase domains (Cox, 2007; Smith & Tsai, 2007; Hertweck, 2009). In fungi, the different catalytic activities often work in an iterative manner (fungal

type I) and it is generally difficult to predict the exact product formed by a given PKS from its sequence alone (Keller *et al.*, 2005). Product prediction is further complicated by the fact that the resulting polyketide structure may be decorated by tailoring enzymes. Such genes are often physically associated with the PKS gene in a gene cluster allowing for coordinated regulation (Schümann & Hertweck, 2006). The fact that natural products may be of mixed biosynthetic origin, combining elements such as polyketides with terpenes (meroterpenoids) and/or nonribosomal peptide units, adds to the complexity (Chang *et al.*, 2009; Geris & Simpson, 2009; Hertweck, 2009; Scherlach *et al.*, 2010).

As a consequence of their bioactivity, societal importance and also the prospect of reprogramming the biosynthetic machinery for drug development (Cox, 2007), there is tremendous interest in the discovery and understanding of fungal polyketide biosynthesis. The availability of full genome sequences of a number of filamentous fungi has provided a means to address the discovery of polyketides because the PKS genes are large and contain several conserved protein domains. Importantly, analysis of the

genomic sequences from filamentous fungi (including *Aspergillus nidulans*, teleomorph, *Emericella nidulans*) predict numbers of individual PKS genes that exceeds significantly the number of polyketides that these fungi are known to produce (Galagan *et al.*, 2005). In fact, the genome of *A. nidulans* (Galagan *et al.*, 2005) appears to contain as many as 32 individual PKS genes (Nierman *et al.*, 2005; Szewczyk *et al.*, 2008; von Döhren, 2009), but until now only nine genes have been linked to eight polyketides (Yamazaki & Maebayas, 1982; Bergmann *et al.*, 2007; Chiang *et al.*, 2008; Szewczyk *et al.*, 2008; Bok *et al.*, 2009; Chiang *et al.*, 2009; Schroeckh *et al.*, 2009) (see Supporting Information, Fig. S1). This may reflect that many polyketides are either produced in small amounts, under special conditions, or in developmental stages that are rarely observed under laboratory conditions.

Towards a more complete genetic mapping of the secondary metabolism in *A. nidulans*, we first cultivated a reference strain on an array of different growth media to uncover polyketides that were not previously linked to a gene cluster. This analysis revealed several compounds, including austinols, violaceols, arugosins and prenylated xanthenes. Next, genetic links to these compounds were established by constructing and screening an *A. nidulans* mutant library containing individual deletions of 32 putative PKS genes.

Materials and methods

Strains

The *A. nidulans* strain IBT29539 (*argB2*, *pyrG89*, *veA1*, *nkuAΔ*) (Nielsen *et al.*, 2008) was used as the reference strain and for deletion-strain constructions. *Escherichia coli* strain DH5α was used for cloning.

Media

Fungal minimal medium (MM) was as described in Cove (1966), but with 1% glucose, 10 mM NaNO₃ and 2% agar. Medium for *alcA* promoter induction consisted of MM supplemented with 100 mM L-threonine and 100 mM glycerol as carbon source instead of 1% glucose. Polyketide screening media variants CYA, CYAs, YES and RT were prepared as per Frisvad & Samson (2004). CY20 medium consisted of CYA with 170 g sucrose; RTO contained RT with 30 g organic oat meal; and YE was made as YES but without sucrose. All media variants were supplemented with 10 mM uridine, 10 mM uracil and/or 4 mM L-arginine where appropriate.

Construction of *A. nidulans* gene deletion library

Individual PKS gene deletions were carried out as described previously (Nielsen *et al.*, 2006), except that the targeting fragments were assembled using Gateway technology (Hartley

et al., 2000) (see Table S1 for PCR oligonucleotide and Fig. S2 for an overview of the procedure). The *A. nidulans* transformants were streak purified and rigorously screened through three complementing diagnostic PCRs. Subsequently, the *Aspergillus fumigatus pyrG* marker was eliminated from all strains by selecting on 5-fluoroorotic acid medium before final verification by two additional complementing diagnostic PCRs (see Fig. S3 and Table S2). All strains have been deposited in the IBT strain collection, DTU, (<http://www.fbd.dtu.dk/straincollection/>).

Construction of the *ausA*-S1660A strain

The amino acid substitution of serine to alanine, S1660A, in *ausA* (AN8383) was created by USER fusion (Geu-Flores *et al.*, 2007) according to the method described by Hansen *et al.* (2011). The allele was transferred to IBT29539 and the *pyrG* marker was eliminated by direct repeat recombination, creating strain IBT31032 containing only the desired point mutation. The strain was verified to contain the *ausA*-S1660A allele by sequencing (StarSEQ, Germany). See Table S3 for primer details.

Ectopic integration of *ausA* into IS1

The gene, *ausA*, was PCR amplified by USER fusion (Geu-Flores *et al.*, 2007) and inserted into both pU1111-1 and pU1211-1 (Hansen *et al.*, 2011) creating pDH23 (*gpdA* promoter) and pDH24 (*alcA* promoter), respectively. For both plasmids, the gene-targeting substrate was excised by NotI digestion and transformed into IBT28738 using *A. nidulans argB* as a selectable marker. Transformants were streak purified and verified for correct integration into the IS1 site (Hansen *et al.*, 2011) by two complementing diagnostic PCRs.

Chemical analysis of the mutants

Strains were inoculated as three point stabs on solid media and incubated for 7 days at 37 °C in the dark. Metabolite extraction was performed according to the micro extraction procedure (Smedsgaard, 1997). Extracts were analyzed by two methods: (1) Ultra-high performance liquid chromatography-diode array detection (UHPLC-DAD) analyses using a Dionex RSLC Ultimate 3000 (Dionex, Sunnyvale, CA) equipped with a diode-array detector. Separation of 1 µL extract was obtained on a Kinetex C₁₈ column (150 × 2.1 mm, 2.6 µm; Phenomenex, Torrance, CA) at 60 °C using a linear water–acetonitrile gradient starting from 15% CH₃CN to 100% (50 ppm trifluoroacetic acid) over 7 min at a flow rate of 0.8 mL min⁻¹. (2) Exact mass, HPLC-DAD-HRMS, was performed on a 5 cm × 3 µm, Luna C18(2) column (Phenomenex) using a water–acetonitrile gradient from 15% CH₃CN to 100% over 20 min (20 mM formic acid). LC-DAD-MS analysis was performed on a LCT oaTOF mass spectrometer (Micromass,

Manchester, UK) as in Nielsen & Smedsgaard (2003) and Nielsen *et al.* (2009).

Isolation and structure elucidation of secondary metabolites

3,5-Dimethylorsellinic acid and dehydroaustinol were purified from large-scale ethyl acetate extracts prepared from 100 MM agar plates after 4 days' cultivation in darkness at 37 °C. The compounds were purified using a 10 × 250 mm Phenomenex pentafluorophenyl column (5 µm particles) with a water–acetonitrile gradient from 15% to 100% CH₃CN in 20 min using a flow of 5 mL min⁻¹. Arugosin A was isolated from an ethyl acetate extract of the reference strain grown on 200 CYAs agar plates using a Waters 19 × 300 mm C18 Delta Pak column (15 µm particles), gradient from 80% to 90% CH₃CN in 10 min, and a flow of 30 mL min⁻¹. The NMR spectra were acquired on a Varian Unity Inova 500 MHz spectrometer using standard pulse

sequences. Additional details about the compound identification can be found in the supporting information.

Results and discussion

Selecting media that support secondary metabolite production in *A. nidulans*

The principle of using different media and/or incubation conditions for fungal secondary metabolite production has often been promoted (Oxford *et al.*, 1935; Davis *et al.*, 1966; Pitt *et al.*, 1983; Bode *et al.*, 2002; Scherlach & Hertweck, 2006). Based on our previous experiences (Frisvad, 1981; Frisvad & Filtenborg, 1983; Filtenborg *et al.*, 1990; Frisvad *et al.*, 2007), eight different media, CYA, CYAs, CY20, MM, RT, RTO, YE and YES, were initially selected for the analysis (Fig. 1a). HPLC analyses revealed a large number of different secondary metabolites produced by the *A. nidulans* reference strain on CYA, CYAs, CY20, RT, RTO and YES (Fig. 1b) and these metabolites served as a source for further investigation.

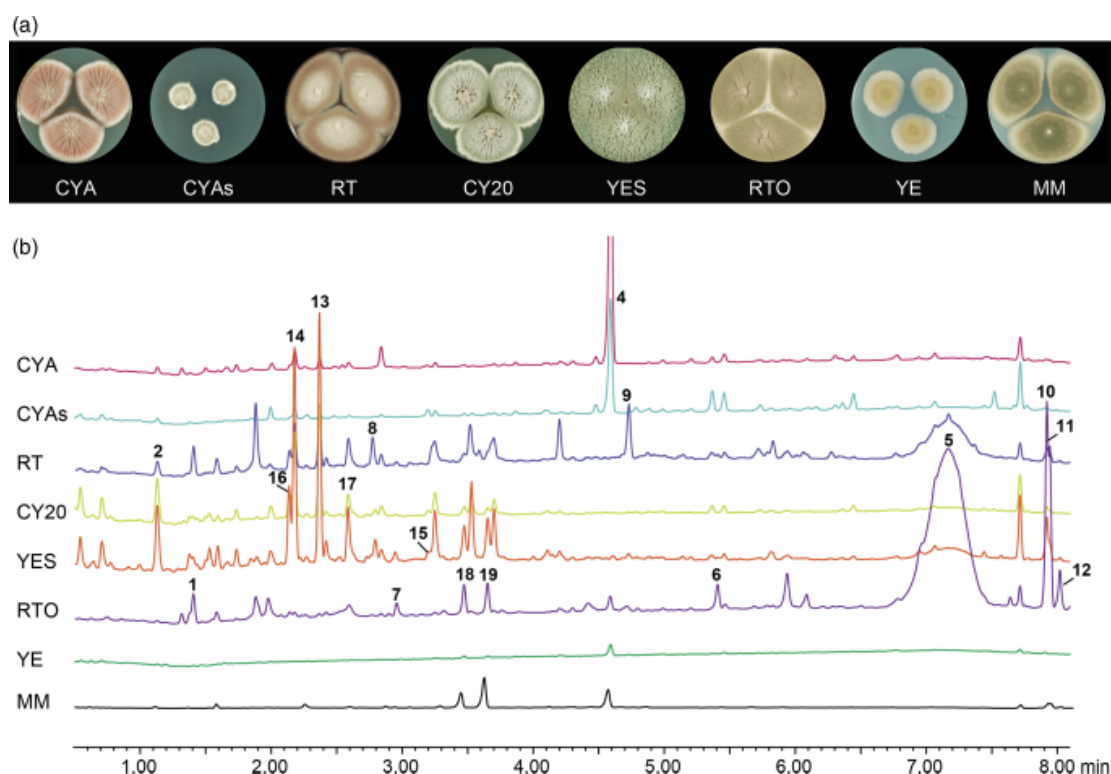


Fig. 1. (a) The appearance of the *Aspergillus nidulans* reference strain, IBT29539, cultivated on different media. (b) UV-chromatograms (210 nm) of extracts of the reference strain provides an excellent demonstration of the profound effect different media compositions have on production of secondary metabolites. From the bottom: minimal media (MM), yeast extract (YE), Raulin-Thom oatmeal (RTO), yeast extract sucrose (YES), Czapek yeast salt (CY20), Raulin-Thom (RT), Czapek yeast agar salt (CYAs), and Czapek yeast agar (CYA). The metabolite profiles on MM and YE media were considered to be redundant and these two media were therefore not used for further screening. Numbers above peaks correspond to compound identity as follows: monodictyphenone (1), orsellinic acid (2), sterigmatocystin (4), arugosin A (5), arugosin H (6), 2- ω -dihydroxyemodin (7), ω -hydroxyemodin (8), emodin (9), emericellin (10), shamixanthone (11), epi-shamixanthone (12), violaceol I (13), violaceol II (14), lecanoric acid (15), F9775A (16), F9775B (17), austinol (18) and dehydroaustinol (19).

Construction and initial characterization of a genome-wide PKS gene deletion library

To investigate whether any of the compounds observed in Fig. 1 could be genetically linked to a PKS gene, we decided to take a global approach and individually deleted all 32 (putative and known) PKS genes in the *A. nidulans* genome (see Fig. S4), as defined from the annotation of the genome databases at the Broad Institute of Harvard and MIT and the

Aspergillus Genome Database at Stanford (Arnaud *et al.*, 2010). All genes were individually deleted by replacing the entire ORFs using gene-targeting substrates based on the *pyrG* marker from *A. fumigatus* for selection. Before analyzing the deletion mutant strains, the *pyrG* marker was excised by direct repeat recombination (Nielsen *et al.*, 2006) in each case. This was carried out to ensure that the analyses of individual mutant strains were comparable to and not influenced by differences in the primary metabolism due to gene cluster-specific expression levels of the *pyrG* marker. All 32 deletion mutant strains (see Table S4) were viable and able to sporulate, showing that none of the 32 genes are essential for growth and that no polyketide product is essential for conidiation. As expected, the one strain carrying the *wAΔ* mutation formed white conidiospores as it fails to produce the naphthopyrone, YWA1, the precursor of green conidial pigment (Watanabe, 1998; Watanabe *et al.*, 1999).

In addition to *wA*, eight additional PKS genes have previously been linked to metabolites. In our analysis, key compounds representing four of these gene clusters could be detected: monodictyphenone (**1**) (observed on RTO, YES and CY20), orsellinic acid (**2**) (observed on YES, CY20, RT, CYAs and CYA), emericellamide (A) (**3**) (observed on all media) and sterigmatocystin (**4**) (observed on RTO, CYAs and CYA). To verify the previously published gene links to these compounds, we individually compared the metabolic profiles of the reference strain to the corresponding profiles obtained with the single PKS gene deletion mutant strains. In agreement with previous analyses, these four compounds

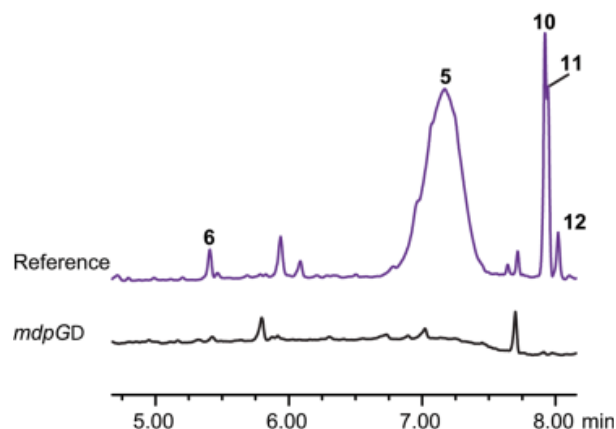


Fig. 2. The late portion of the UV-chromatogram (210 nm) of the *Aspergillus nidulans* reference strain and *mdpGΔ* on RTO medium. The deletion abolishes formation of arugosin A (**5**), arugosin H (**6**), emericellin (**10**), shamixanthone (**11**) and epi-shamixanthone (**12**) and a few unidentified peaks.

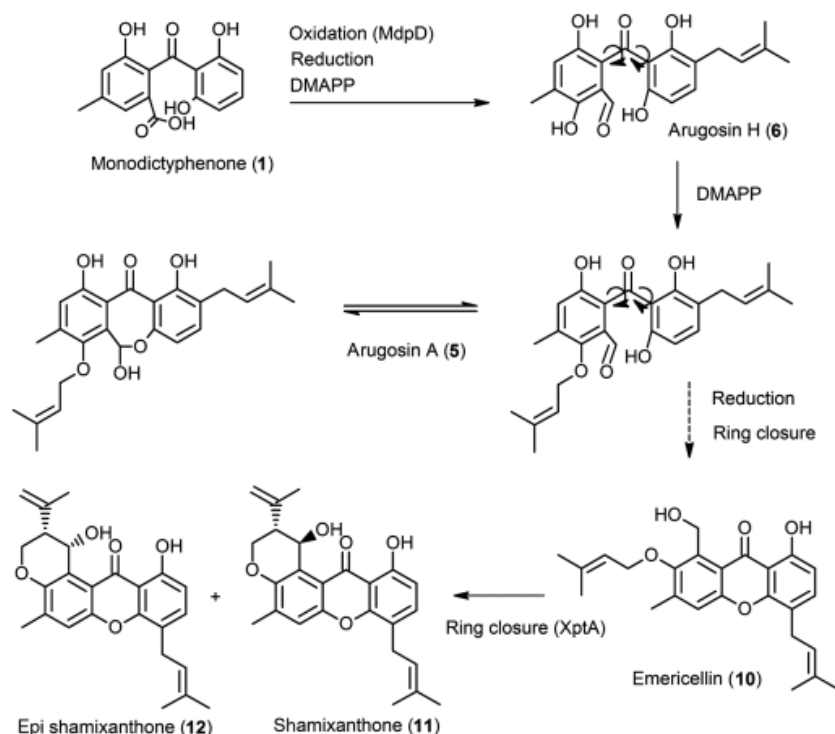


Fig. 3. Proposed biosynthetic route of the arogusins and shamixanthones. As seen in Fig. 2, arugosin A is produced in large amounts relative to the expected end products: emericellin (**10**) and shamixanthones (**11** and **12**). DMAPP, dimethylallylpyrophosphate. MdpG and XptA are enzymatic activities as proposed by Sanchez *et al.* (2011).

disappeared in *mdpG* (Bok *et al.*, 2009), *orsA* (Schroeckh *et al.*, 2009), *easB* (Chiang *et al.*, 2008) and *stcA* (Yu & Leonard, 1995) deletion strains of our library (Fig. S5). Compounds resulting from the remaining four PKS genes were identified by activating the gene clusters by controlled expression of the transcription factor gene in the cluster (Bergmann *et al.*, 2007; Chiang *et al.*, 2009) or by deleting *sumO* that influences regulation of biological processes at many different levels (Szewczyk *et al.*, 2008). Expression from these clusters is apparently not triggered by growth on any of our media, and natural conditions provoking their activation remain to be discovered.

Next, we performed a comparison of the metabolite profiles from the 32 deletion mutants with those obtained with the reference strain with the aim of uncovering novel genetic links between PKS genes and polyketides. The most significant changes are described below.

mdpG is involved in the biosynthesis of arugosins

First we focused our attention on the most prominent compound produced on RTO, YES, CY20 and RT media, which eluted as a broad peak around 7.2 min. This compound completely disappeared in the *mdpG* Δ strain (Fig. 2 and Fig. S6). Moreover, it had a characteristic UV spectrum in the UHPLC analysis indicating an arugosin-like metabolite. MS analysis indicated the compound to be arugosin A (m/z 425 a.m.u. for $[M+H]^+$), which to our knowledge has not been reported before from *A. nidulans*. We therefore decided to confirm the structure of this compound (5). A large-scale extraction was performed and the metabolite was purified. The NMR data in dimethyl sulfoxide are in agreement with the literature (Kawahara *et al.*, 1988) for the hemiacetal form of arugosin A except that the equilibrium was shifted completely to the open form (Fig. 3). In methanol, the NMR data showed that the compound exists in equilibrium between the closed and open ring form (data not shown), explaining the broad peak observed in Fig. 2. A minor peak could be assigned as a mono-prenylated arugosin as $[M+H]^+$ at m/z 357 a.m.u. The MS data of this compound did not indicate loss of a prenyl moiety, suggesting that it is arugosin H (6), a likely immediate precursor of arugosin A (Fig. 3). Hence, our data show that *mdpG*, which is known for its role in formation of monodictyphenone, is also involved in formation of arugosins.

It is not unusual that one PKS gene cluster is responsible for the biosynthesis of a family of structurally similar compounds (Walsch, 2002; Kroken *et al.*, 2003; Frisvad *et al.*, 2004; Amoutzias *et al.*, 2008). In the original analysis of the *mdpG* gene cluster, it was activated due to remodeling of the chromatin landscape, which occurs in a *cclA* deletion strain (Chiang *et al.*, 2010). That study genetically linked the

mdpG gene cluster to eight emodin analogues, including several aminated species, which were detected and tentatively identified. In our analyses, we also detected several emodins including 2- ω -dihydroxyemodin (7), ω -hydroxyemodin (8) and emodin (9), as well as the more apolar compounds emericellin (10), shamixanthone (11) and epishamixanthone (12) (Fig. 1 and Fig. S7). Like in the original study, all emodins disappear in our *mdpG* Δ strain.

Recently, it was demonstrated that the polyketide part of prenylated xanthenes also could be coupled to *mdpG* (Sanchez *et al.*, 2011). Our finding that *mdpG* is involved in formation of arugosins indicates that these compounds serve as intermediates in the conversion of monodictyphenone into xanthenes, Fig. 3. In agreement with this, previous studies have reported arugosins to be precursors for emericellin (10) and shamixanthenes (11) and (12) (Ahmed *et al.*, 1992; Kralj *et al.*, 2006; Márquez-Fernández *et al.*, 2007), but have not established a genetic link to *mdpG*.

AN7903 and *orsA* are involved in violaceol production

Our reference strain produces the antibiotic violaceol I (13) and II (14), in significant amounts (Fig. 4 and Fig. S8). These two diphenyl ethers have been identified in *Emericella*

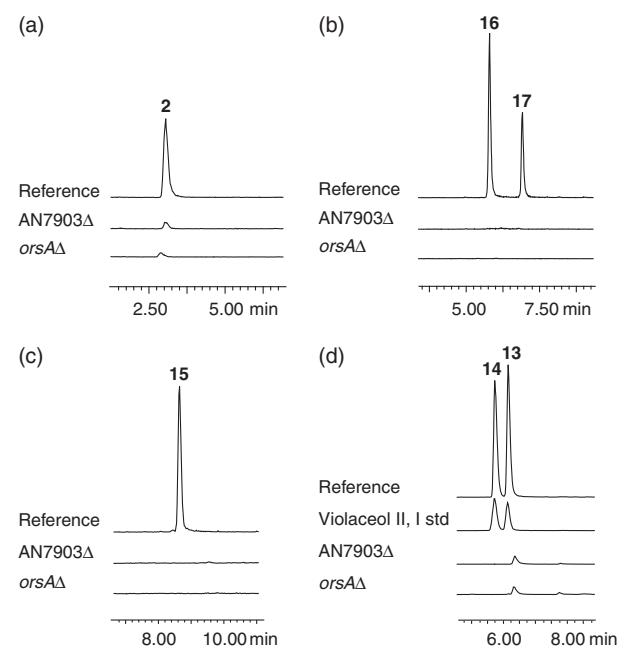


Fig. 4. Extracted ion chromatograms of metabolic extracts from the reference strain and the two mutants *orsA* Δ and AN7903 Δ after cultivation on YES medium. Compounds shown are (a) m/z 167 $[M-H]^-$, orsellinic acid (2), (b) m/z 397 $[M+H]^+$, F9775A and B (16 and 17, respectively), (c) m/z 317 $[M-H]^-$, lecanoric acid (15) and (d) m/z 261 $[M-H]^-$, violaceol I and II (13 and 14). An extracted ion chromatogram of a violaceol I and II in house standard is included in (d).

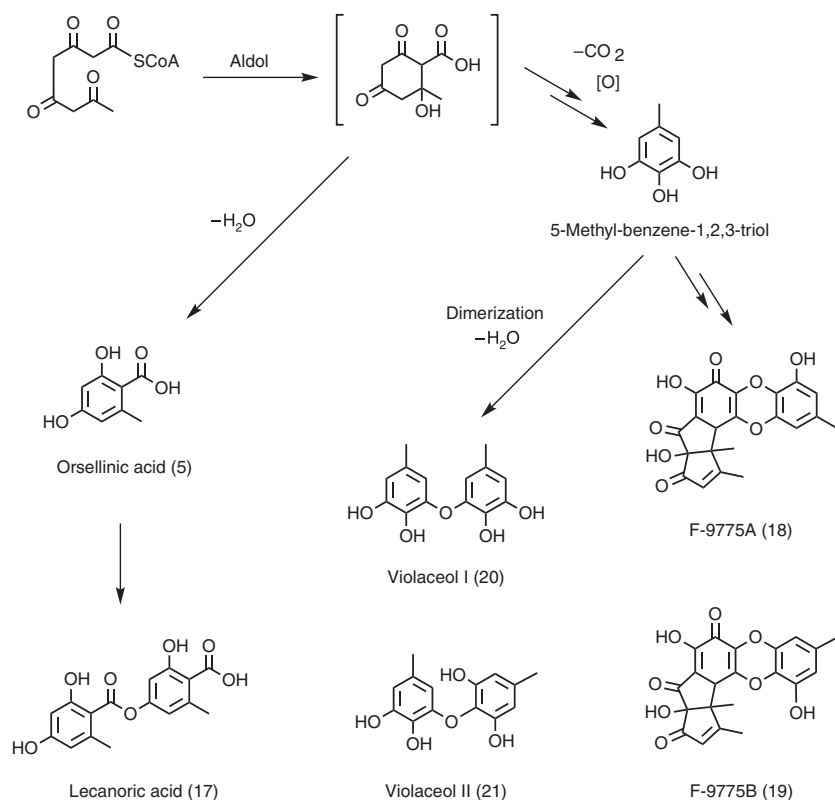


Fig. 5. Proposed biosynthetic pathways related to the *orsA* gene cluster. A common C₈ aldol intermediate, presented in the bracket, can by loss of water and enolization lead to orsellinic acid and by further esterification to lecanoric acid, or be decarboxylated and further oxidized to yield the C₇ compound 5-methyl-benzene-1,2,3-triol. This compound is the precursor of the two diphenyl ethers violaceol I and II and the pigments F-9775A and B.

violacea, *Aspergillus sydowii* and *Aspergillus funiculosus* (Taniguchi *et al.*, 1978; Yamazaki & Maebayas, 1982) and recently also in *A. nidulans* (Nahlik *et al.*, 2010). Our analysis now links the biosynthesis of the two violaceols to *orsA* as they disappear in our *orsA*Δ strain. It has previously been shown that *orsA* (AN7909) is involved in the formation of orsellinic acid (2), lecanoric acid (15), the two colored compounds F-9775A (16) and F-9775B (17), orcinol, diorcinol, gerfeldin and deoxy-gerfeldin. (Schroeckh *et al.*, 2009; Sanchez *et al.*, 2010). Our analysis confirms the link between orsellinic acid, lecanoric acid, diorcinol, F-9775A, F-9775B to *orsA* as these compounds are missing in the *orsA*Δ strain. However, we have not been able to detect the gerfeldins in any of our strains, and apparently our conditions favor violaceol and not gerfeldin formation.

The violaceols are formed by dimerization of two C₇ monomers of 5-methylbenzene-1,2,3-triol, a compound that we could tentatively detect as [M-H][−] at *m/z* 139 in cultivation extracts. The C₇ backbone of 5-methylbenzene-1,2,3-triol, may conceivably be formed by decarboxylation of a C₈ aldol intermediate as suggested by Turner 40 years ago (Turner, 1971) (Fig. 5). This C₈ intermediate also serves as a branch point towards orsellinic acid.

Interestingly, the same compounds that disappear in the *orsA*Δ strain also disappear in AN7903Δ, a strain missing a PKS gene separated from *orsA* by only ~20 kb (Fig. 4). This

result does not contradict the original assignment of *orsA* as the PKS gene responsible for production of orsellinic acid. Although the enzymes encoded by the two genes are predicted to share many of the same functional domains, AN7903 is larger by around 500 amino acid residues and contains a methyl-transferase domain, which is not required for orsellinic acid production. Moreover, we note that Schroeckh *et al.* (2009) observed that both AN7903 and *orsA* were upregulated when orsellinic acid was induced by co-cultivation with *Streptomyces hygroscopicus*, indicating cross-talk between the two clusters. Surprisingly, what appear to be trace amounts of orsellinic acid can be detected as *m/z* 167 [M-H][−] in both the AN7903Δ and the *orsA*Δ strains (Fig. 4). The source of this residual orsellinic acid remains elusive, but it could possibly stem from unmethylated byproducts from the PKS, AN8383, that produces 3,5-dimethylorsellinic acid, see below.

AN8383 is responsible for austinol and dehydroaustinol biosynthesis

Interestingly, production of austinol (18) and dehydroaustinol (19) was observed in the reference strain on several media (Fig. 1). Despite the fact that the production of these compounds is known from *A. nidulans* (Szewczyk *et al.*, 2008), they have not yet been assigned to a specific gene.

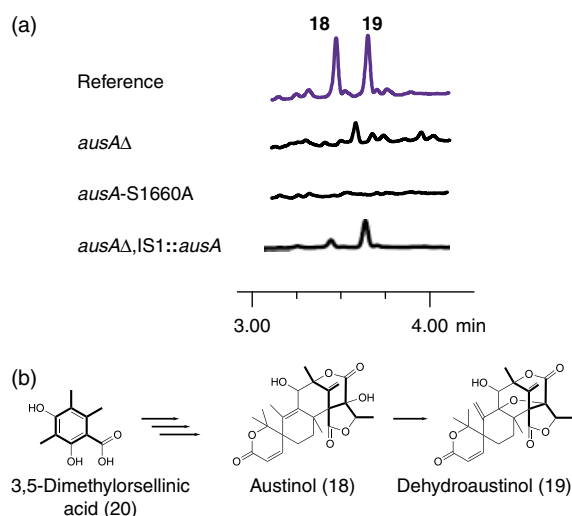


Fig. 6. (a) UV-chromatograms (210 nm) showing the presence of austinol (**18**) and dehydroaustinol (**19**) in the reference strain, *ausAΔ* and *ausA-S1660A* after growth on RTO medium. The lower trace (*ausAΔ, IS1::ausA*) is the *ausAΔ* mutant expressing *ausA* under control of the *gpdA* promoter from IS1 (IBT31281, cultured on MM). (b) Incorporation of 3,5-orsellinic acid (**20**) into the meroterpenoid austinol. The carbon skeleton of orsellinic acid is presented in bold.

Only the AN8383Δ strain failed to produce the two austinols on all the media, which triggered austinol production in the reference strain (Fig. 6a). This, phenotype could be rescued by inserting the structural gene of AN8383 under the control of the *gdpA* promoter into an ectopic locus, IS1 (Hansen *et al.*, 2011) (Fig. 6a). Moreover, a point mutant strain AN8383-S1660A also failed to produce austinols on these six media (Fig. 6a). In this strain, the DSL motif of the AN8383 PKS has been mutated to DAL, preventing the phosphopantetheine moiety of coenzyme A to attach to the acyl carrier protein domain of the PKS, thus disrupting polyketide synthesis (Evans *et al.*, 2008). The lack of austinols can thus be linked directly to an AN8383-encoded function rather than to silencing of another gene caused by chromatin changes provoked by the AN8383 deletion.

To confirm the role of AN8383 in austinol production, we constructed a new strain that expresses the AN8383 ORF controlled by the inducible *alcA* promoter from the ectopic locus, IS1 (Hansen *et al.*, 2011). On inductive medium, the subsequent LC-MS analysis showed a large novel peak eluting at ca. 6 min (see Fig. S9). The corresponding compound was purified and the structure was elucidated by NMR (Fig. S10), identifying 3,5-dimethylorsellinic acid (**20**), which is in good agreement with the route of synthesis previously proposed for austinol (Fig. 6b; Geris & Simpson, 2009). In a parallel analysis using a strain expressing AN8383 under the control of the constitutive *gpdA* promoter we obtained the same result (data not shown).

Together, the results strongly indicate that AN8383 encodes a PKS producing 3,5-dimethylorsellinic acid and that this compound serves as the first intermediate in the complex biosynthesis of austinol and dehydroaustinol that also involves a yet unidentified prenyl transferase(s). Based on these results, we have named the locus AN8383, *ausA*.

Concluding remarks

In the present study, we constructed a genome-wide PKS deletion library, which we screened for effects on polyketide production on a variety of media. This analysis has provided novel links between PKS genes and polyketide products demonstrating the strength of this approach. Many PKS genes still remain to be functionally connected to products, as many gene clusters have not yet been activated. As the repertoire of tools and methods to induce gene expression is rapidly increasing, new polyketide compounds will likely soon be uncovered in *A. nidulans*. To this end, the genome-wide PKS gene deletion library presented here will undoubtedly serve as a useful resource.

Acknowledgements

This work was supported by the Danish Research Agency for Technology and Production, grant # 09-064967. We thank Lisette Knoth-Nielsen for her dedicated and valuable technical assistance in the laboratory. In addition, we thank Rasmus John Normand Frandsen for suggestions and critical reading of the manuscript.

References

- Ahmed SA, Bardshiri E, McIntyre CR & Simpson TJ (1992) Biosynthetic studies on Tajixanthone and Shamixanthone, Polyketide Hemiterpenoid Metabolites of *Aspergillus varicolor*. *Aust J Chem* **45**: 249–274.
- Amoutzias GD, Van de Peer Y & Mossialos D (2008) Evolution and taxonomic distribution of nonribosomal peptide and polyketide synthases. *Future Microbiol* **3**: 361–370.
- Arnaud MB, Chibucos MC, Costanzo MC *et al.* (2010) The *Aspergillus* genome database, a curated comparative genomics resource for gene, protein and sequence information for the *Aspergillus* research community. *Nucleic Acids Res* **38**: D420–D427.
- Bentley R (2000) Mycophenolic acid: a one hundred year Odyssey from antibiotic to immunosuppressant. *Chem Rev* **100**: 3801–3825.
- Bérday J (2005) Bioactive microbial metabolites. *J Antibiot (Tokyo)* **58**: 1–26.
- Bergmann S, Schumann J, Scherlach K, Lange C, Brakhage AA & Hertweck C (2007) Genomics-driven discovery of PKS-NRPS hybrid metabolites from *Aspergillus nidulans*. *Nat Chem Biol* **3**: 213–217.

- Bode HB, Bethe B, Höfs R & Zeeck A (2002) Big effects from small changes: possible ways to explore nature's chemical diversity. *ChemBiochem* **3**: 619–627.
- Bok JW, Chiang YM, Szeewczyk E *et al.* (2009) Chromatin-level regulation of biosynthetic gene clusters. *Nat Chem Biol* **5**: 462–464.
- Chang P-K, Ehrlich KC & Fujii I (2009) Cyclopiazonic acid biosynthesis of *Aspergillus flavus* and *Aspergillus oryzae*. *Toxins* **1**: 74–99.
- Chiang Y-M, Szeewczyk E, Nayak T *et al.* (2008) Molecular genetic mining of the *Aspergillus* secondary metabolome: discovery of the emericellamide biosynthetic pathway. *Chem Biol* **15**: 527–532.
- Chiang Y-M, Szeewczyk E, Davidson AD, Keller N, Oakley BR & Wang CCC (2009) A gene cluster containing two fungal polyketide synthases encodes the biosynthetic pathway for a polyketide, asperfuranone, in *Aspergillus nidulans*. *J Am Chem Soc* **131**: 2965–2970.
- Chiang Y-M, Szeewczyk E, Davidson AD, Entwistle R, Keller N, Wang CCC & Oakley BR (2010) Characterization of the *Aspergillus nidulans* monodictyphenone gene cluster. *Appl Environ Microb* **76**: 2067–2074.
- Chooi Y-H, Cacho R & Tang Y (2010) Identification of the viridicatumtoxin and griseofulvin gene clusters from *Penicillium aethiopicum*. *Chem Biol* **17**: 483–494.
- Cove DJ (1966) The induction and repression of nitrate reductase in the fungus *Aspergillus nidulans*. *Biochim Biophys Acta* **113**: 51–56.
- Cox RJ (2007) Polyketides, proteins and genes in fungi: programmed nano-machines begin to reveal their secrets. *Org Biomol Chem* **5**: 2010–2026.
- Davis ND, Diener UL & Eldridge DW (1966) Production of aflatoxins B1 and G1 by *Aspergillus flavus* in a semisynthetic medium. *Appl Microbiol* **14**: 378–380.
- Evans SE, Williams C, Arthur CJ, Burston SG, Simpson TJ, Crosby J & Crump MP (2008) An ACP structural switch: conformational differences between the apo and holo forms of the actinorhodin polyketide synthase acyl carrier protein. *ChemBioChem* **9**: 2424–2432.
- Filtenborg O, Frisvad JC & Thrane U (1990) The significance of yeast extract composition on metabolite production in *Penicillium*. *Modern Concepts in Penicillium and Aspergillus Classification* (Samson RA & Pitt JI, eds), pp. 433–440. Plenum Press, New York.
- Frisvad JC (1981) Physiological criteria and mycotoxin production as aids in identification of common asymmetric penicillia. *Appl Environ Microb* **41**: 568–579.
- Frisvad JC & Filtenborg O (1983) Classification of terverticillate penicillia based on profiles of mycotoxins and other secondary metabolites. *Appl Environ Microb* **46**: 1301–1310.
- Frisvad JC & Samson RA (2004) Polyphasic taxonomy of *Penicillium* subgenus *Penicillium*. A guide to identification of food and air-borne terverticillate Penicillia and their mycotoxins. *Stud Mycol* **49**: 1–173.
- Frisvad JC, Smedsgaard J, Larsen TO & Samson RA (2004) Mycotoxins, drugs and other extrolites produced by species in *Penicillium* subgenus *Penicillium*. *Stud Mycol* **49**: 201–241.
- Frisvad JC, Smedsgaard J, Samson RA, Larsen TO & Thrane U (2007) Fumonisin B(2) production by *Aspergillus niger*. *J Agr Food Chem* **55**: 9727–9732.
- Galagan JE, Calvo SE, Cuomo C *et al.* (2005) Sequencing of *Aspergillus nidulans* and comparative analysis with *A. fumigatus* and *A. oryzae*. *Nature* **438**: 1105–1115.
- Geris R & Simpson TJ (2009) Meroterpenoids produced by fungi. *Nat Prod Rep* **26**: 1063–1094.
- Geu-Flores F, Nour-Eldin HH, Nielsen MT & Halkier BA (2007) USER fusion: a rapid and efficient method for simultaneous fusion and cloning of multiple PCR products. *Nucleic Acids Res* **35**: e55.
- Hansen BG, Salomonsen B, Nielsen MT, Nielsen JB, Hansen NB, Nielsen KF, Regueira TB, Nielsen J, Patil KR & Mortensen UH (2011) A versatile gene expression and characterization system for *Aspergillus*: heterologous expression of the gene encoding the polyketide synthase from the mycophenolic acid gene cluster from *Penicillium brevicompactum* as a case study. *Appl Environ Microb* **77**: 3044–3051.
- Hartley JL, Temple GF & Brasch MA (2000) DNA cloning using *in vitro* site-specific recombination. *Genome Res* **10**: 1788–1795.
- Hertweck C (2009) The biosynthetic logic of polyketide diversity. *Angew Chem Int Edit* **48**: 4688–4716.
- Kawahara N, Nozawa K, Nakajima S & Kawai K (1988) Studies on fungal products. Part 15. Isolation and structure determination of arugosin E from *Aspergillus silvaticus* and cycloisoemicellin from *Emericella striata*. *J Chem Soc Perk Trans I* **1988**: 907–911.
- Keller N, Turner G & Bennett J (2005) Fungal secondary metabolism – from biochemistry to genomics. *Nat Rev Microbiol* **3**: 937–947.
- Kralj A, Kehraus S, Krick A, Eguereva E, Kelter G, Maurer M, Wortmann A, Fiebig HH & König GM (2006) Arugosins G and H: prenylated polyketides from the marine-derived fungus *Emericella nidulans* var. *acristata*. *J Nat Prod* **69**: 995–1000.
- Kroken S, Glass NL, Taylor JW, Yoder OC & Turgeon BG (2003) Phylogenomic analysis of type I polyketide synthase genes in pathogenic and saprobic ascomycetes. *P Nat Acad Sci USA* **100**: 15670–15675.
- Lai L-ST, Tsai T-H, Wang TC & Cheng T-Y (2005) The influence of culturing environments on lovastatin production by *Aspergillus terreus*. *Enzyme Microb Tech* **36**: 737–748.
- Månsson M, Klejnstrup ML, Phipps RK, Nielsen KF, Frisvad JC, Gotfredsen CH & Larsen TO (2010) Isolation and NMR characterization of fumonisin B2 and a new fumonisin B6 from *Aspergillus niger*. *J Agr Food Chem* **58**: 949–953.
- Márquez-Fernández O, Trigos A, Ramos-Balderas JL, Viniegra-González G, Deising HB & Aguirre J (2007) Phosphopantetheinyl transferase CfwA/NpgA is required for *Aspergillus nidulans* secondary metabolism and asexual development. *Eukaryot cell* **6**: 710–720.

- Nahlik K, Dumkow M, Bayram O *et al.* (2010) The COP9 signalosome mediates transcriptional and metabolic response to hormones, oxidative stress protection and cell wall rearrangement during fungal development. *Mol Microbiol* **78**: 964–979.
- Newman DJ & Cragg GM (2007) Natural products as sources of new drugs over the last 25 years. *J Nat Prod* **70**: 461–477.
- Nielsen JB, Nielsen ML & Mortensen UH (2008) Transient disruption of non-homologous end-joining facilitates targeted genome manipulations in the filamentous fungus *Aspergillus nidulans*. *Fungal Genet Biol* **45**: 165–170.
- Nielsen KF & Smedsgaard J (2003) Fungal metabolite screening: database of 474 mycotoxins and fungal metabolites for dereplication by standardised liquid chromatography-UV-mass spectrometry methodology. *J Chromatogr A* **1002**: 111–136. ISI: 000183799200011.
- Nielsen KF, Mogensen JM, Johansen M, Larsen TO & Frisvad JC (2009) Review of secondary metabolites and mycotoxins from the *Aspergillus niger* group. *Anal Bioanal Chem* **395**: 1225–1242.
- Nielsen ML, Albertsen L, Lettier G, Nielsen JB & Mortensen UH (2006) Efficient PCR-based gene targeting with a recyclable marker for *Aspergillus nidulans*. *Fungal Genet Biol* **43**: 54–64.
- Nierman WC, Pain A, Anderson MJ *et al.* (2005) Genomic sequence of the pathogenic and allergenic filamentous fungus *Aspergillus fumigatus*. *Nature* **438**: 1151–1156.
- Oxford AE, Raistrick H & Simonart P (1935) Studies in the biochemistry of microorganisms. XLIV. Fulvic acid, a new crystalline yellow pigment, a metabolic product of *P. griseofulvum* Dierckx, *P. flexuosum* Dale and *P. brefeldianum* Dodge. *Biochem J* **29**: 1102–1115.
- Pitt JI, Hocking AD & Glenn DR (1983) An improved medium for the detection of *Aspergillus flavus* and *A. parasiticus*. *J Appl Bacteriol* **54**: 109–114.
- Sanchez JF, Chiang YM, Szewczyk E, Davidson AD, Ahuja M, Elizabeth Oakley C, Woo Bok J, Keller N, Oakley BR & Wang CC (2010) Molecular genetic analysis of the orsellinic acid/F9775 gene cluster of *Aspergillus nidulans*. *Mol Biosyst* **6**: 587–593.
- Sanchez JF, Entwistle R, Hung J, Yaegashi J, Jain S, Chiang Y, Wang C & Oakley B (2011) Genome-based deletion analysis reveals the prenyl xanthone biosynthesis pathway in *Aspergillus nidulans*. *J Am Chem Soc* **133**: 4010–4017.
- Scherlach K & Hertweck C (2006) Discovery of aspoquinolones A–D, prenylated quinoline-2-one alkaloids from *Aspergillus nidulans*, motivated by genome mining. *Org Biomol Chem* **4**: 3517–3520.
- Scherlach K, Schuermann J, Dahse HM & Hertweck C (2010) Aspernidine A and B, prenylated isoindolinone alkaloids from the model fungus *Aspergillus nidulans*. *J Antibiot (Tokyo)* **63**: 375–377.
- Schroekch V, Scherlach K, Nützmann HW, Shelest E, Schmidt-Heck W, Schuermann J, Martin K, Hertweck C & Brakhage AA (2009) Intimate bacterial-fungal interaction triggers biosynthesis of archetypal polyketides in *Aspergillus nidulans*. *P Natl Acad Sci USA* **106**: 14558–14563.
- Schümann J & Hertweck C (2006) Advances in cloning, functional analysis and heterologous expression of fungal polyketide synthase genes. *J Biotechnol* **124**: 690–703.
- Smedsgaard J (1997) Micro-scale extraction procedure for standardized screening of fungal metabolite production in cultures. *J Chromatogr A* **760**: 264–270.
- Smith S & Tsai S-C (2007) The type I fatty acid and polyketide synthases: a tale of two megasynthases. *Nat Prod Rep* **24**: 1041–1072.
- Szewczyk E, Chiang Y-M, Oakley C, Davidson A, Wang C & Oakley B (2008) Identification and characterization of the asperthecin gene cluster of *Aspergillus nidulans*. *Appl Environ Microb* **74**: 7607–7612.
- Taniguchi M, Kaneda N, Shibata K & Kamikawa T (1978) Isolation and biological activity of a self-growth-inhibitor from *Aspergillus sydowi*. *Agr Biol Chem* **42**: 1629–1630.
- Turner WB (1971) *Fungal Metabolites*, Academic Press, London, pp. 1–446.
- von Döhren H (2009) A survey of nonribosomal peptide synthetase (NRPS) genes in *Aspergillus nidulans*. *Fungal Genet Biol* **46** (suppl 1): S45–S52.
- Walsch CT (2002) Combinatorial biosynthesis of antibiotics: challenges and opportunities. *ChemBioChem* **3**: 124–134.
- Watanabe A, Ono Y, Fujii I, Sankawa U, Mayorga ME, Timberlake WE & Ebizuka Y (1998) Product identification of polyketide synthase coded by *Aspergillus nidulans* wA gene. *Tetrahedron Lett* **39**: 7733–7736.
- Watanabe A, Fujii I, Sankawa U, Mayorga M, Timberlake W & Ebizuka Y (1999) Re-identification of *Aspergillus nidulans* wA gene to code for a polyketide synthase of naphthopyrone. *Tetrahedron Lett* **40**: 91–94.
- Yamazaki M & Maebayashi Y (1982) Structure determination of violaceol-I and -II, new fungal metabolites from a strain of *Emericella violacea*. *Chem Pharm Bull (Tokyo)* **30**: 514–518.
- Yu JH & Leonard TJ (1995) Sterigmatocystin biosynthesis in *Aspergillus nidulans* requires a novel type-I polyketide synthase. *J Bacteriol* **177**: 4792–4800.

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Fig. S1. Eight known metabolites that have been linked to specific PKS genes in *Aspergillus nidulans*.

Fig. S2. A graphical illustration of the procedure used to make the gene targeting fragments for the PKS deletions.

Fig. S3. Procedure for diagnostic PCR.

Fig. S4. Chromosome map showing the position of the 32 PKS genes.

Fig. S5. Verification that the polyketide is absent in selected deletion mutants.

Fig. S6. Positive mode extracted ion chromatograms for the *mdpG* Δ strain cultivated on RTO.

Fig. S7. Additional polyketides that were detected in metabolite extracts in this study.

Fig. S8. The mutants *orsA* Δ and AN7903 Δ lack production of violaceols.

Fig. S9. Extracted ion chromatograms of metabolic extracts from the reference and *ausA* Δ strains.

Fig. S10. ^1H NMR spectrum of 3,5-dimethylorsellinic acid in dimethyl sulfoxide- d_6 .

Fig. S11. ^1H NMR spectrum of dehydroaustinol in CDCl_3 .

Fig. S12. ^1H NMR spectrum of arugosin A open form in dimethyl sulfoxide- d_6 .

Table S1. PCR primers for Gateway assembly.

Table S2. Oligonucleotide primers for diagnostic PCR.

Table S3. Additional oligonucleotides used in the study.

Table S4. The constructed *Aspergillus nidulans* strains have been deposited into the IBT Culture Collection.

Please note: Wiley-Blackwell is not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

Chapter 5

Genetics of Polyketide Metabolism in *Aspergillus nidulans*

This chapter comprises the review "Genetics of Polyketide Metabolism in *Aspergillus nidulans*", published in *Metabolites*, of which I am a co-author. The present review describes the characterized genetics of polyketide metabolism in *A. nidulans* until December 2011, and offers an overview of polyketide biosyntheses in *A. nidulans*. This review may be consulted during the following chapters for details.

My primary contribution to this review is the section on austinol and dehydroaustinol, which was linked to the *ausA* gene in the study presented in Chapter 4. As this review was published prior to the studies by Itoh et al. [2012], Lo et al. [2012] and Matsuda et al. [2013], the review does not contain details about austinol and dehydroaustinol biosynthesis, and these will not be discussed in the present thesis.

Review

Genetics of Polyketide Metabolism in *Aspergillus nidulans*

Marie L. Klejstrup¹, Rasmus J. N. Frandsen², Dorte K. Holm², Morten T. Nielsen²,
Uffe H. Mortensen², Thomas O. Larsen¹ and Jakob B. Nielsen^{2,*}

¹ Department of Systems Biology, Center for Microbial Biotechnology, Technical University of Denmark, Søtofts Plads B221, DK-2800 Kgs. Lyngby, Denmark; E-Mails: mark@bio.dtu.dk (M.L.K.); tol@bio.dtu.dk (T.O.L.)

² Department of Systems Biology, Center for Microbial Biotechnology, Technical University of Denmark, Søtofts Plads B223, DK-2800 Kgs. Lyngby, Denmark; E-Mails: rasf@bio.dtu.dk (R.J.N.F.); dmkp@bio.dtu.dk (D.K.H.); motni@bio.dtu.dk (M.T.N.); um@bio.dtu.dk (U.H.M.)

* Author to whom correspondence should be addressed; E-Mail: jbn@bio.dtu.dk; Tel.: +45-4525-2657; Fax: +45-4588-4148.

Received: 1 November 2011; in revised form: 23 December 2011 / Accepted: 17 January 2012 /

Published: 30 January 2012

Abstract: Secondary metabolites are small molecules that show large structural diversity and a broad range of bioactivities. Some metabolites are attractive as drugs or pigments while others act as harmful mycotoxins. Filamentous fungi have the capacity to produce a wide array of secondary metabolites including polyketides. The majority of genes required for production of these metabolites are mostly organized in gene clusters, which often are silent or barely expressed under laboratory conditions, making discovery and analysis difficult. Fortunately, the genome sequences of several filamentous fungi are publicly available, greatly facilitating the establishment of links between genes and metabolites. This review covers the attempts being made to trigger the activation of polyketide metabolism in the fungal model organism *Aspergillus nidulans*. Moreover, it will provide an overview of the pathways where ten polyketide synthase genes have been coupled to polyketide products. Therefore, the proposed biosynthesis of the following metabolites will be presented; naphthopyrone, sterigmatocystin, aspyridones, emericellamides, asperthecin, asperfuranone, monodictyphenone/emodin, orsellinic acid, and the austinols.

Keywords: secondary metabolites; polyketides; polyketide synthases; gene clusters; biosynthesis; *Aspergillus nidulans*

1. Introduction

Aspergillus nidulans, teleomorph *Emericella nidulans*, is one of the most significant biological model systems in the fungal kingdom. This was pioneered by Pontecorvos' [1] work in the middle of the last century, which demonstrated that *A. nidulans*, in addition to the asexual state, also proliferate via sexual and parasexual life cycles, hence, offering an ideal platform for genetic studies. Related species in the genus *Aspergillus* include important industrial cell factories, *A. niger* and *A. oryzae*, species that cause allergic diseases, *A. clavatus*, as well as opportunistic pathogens, such as *A. fumigatus*.

A common feature of aspergilli and filamentous fungi in general is their capacity to produce secondary metabolites (SMs). As opposed to the primary metabolites, SMs are not essential for cellular growth, but provide fungi, as well as bacteria and plants, with a competitive advantage in nature, e.g., by serving as agents for chemical warfare or as signal molecules. Hence, an impressive range of compounds with broad ranging bioactivities has evolved. SMs can be divided into four main chemical classes: Polyketides (PK), terpenoids, shikimic acid derived compounds, and non-ribosomal peptides (NRP). Moreover, hybrid metabolites composed of moieties from different classes are common, as in the meroterpenoids, which are fusions between PKs and terpenes. Hybrid molecules significantly add to the complexity and variety of the fungal metabolomes.

In addition to their likely important ecological roles in their natural biological niches, SMs also have a considerable impact on human life. For instance aflatoxins, ochratoxins, and fumonisins act as mycotoxins by having a detrimental effect on humans and livestock, whereas others are beneficial and serve as food additives, pigments, cholesterol-lowering drugs, immunosuppressants, antibiotics and anticancer agents. The different aspects of SM action and application have spurred a tremendous interest in fungal secondary metabolites, which is further underlined by the fact that around 63% of all small molecule drugs, which reached the market from 1981–2006 were inspired by natural products or derivatives thereof [2].

In filamentous fungi, the competitive race in SM development and the cost of producing and secreting complex compounds have resulted in the evolution of a multifaceted regulation of SM biosynthesis to avoid unnecessary use of resources. This hampers their discovery since production of most SMs is not induced under laboratory conditions. Analysis of full genome sequences of eight different aspergilli have demonstrated that for the majority of genes that putatively encode enzymes for SM production, the product is not known or detected. In this review, we will provide highlights of the use of genome mining, sophisticated molecular biological and chemical tools to trigger the production of SMs from cryptic gene clusters and discuss how these techniques have accelerated our understanding of PK production and regulation in *A. nidulans*.

1.1. Polyketide Biosynthesis in *A. nidulans*

PKs in fungi are synthesized by the use of acyl-CoA units. They act as the general substrates for large multi-domain enzymes named polyketide synthases (PKSs), which resemble eukaryotic fatty-acid synthases (FASs) in domain architecture. PKSs are divided into three types of PKSs based on their catalytic organization, however, only the iterative type I PKS (iPKS) has been reported in *A. nidulans*. The iPKS repeat the use of a single module containing several catalytic domains until the growing

chain of acyl-CoA units block further elongation. For descriptions of PKSs in general, excellent reviews by Crawford [3], Hertweck [4] and Cox [5] can be consulted. The most commonly encountered catalytic activities in fungal PKSs will be addressed as a general introduction to fungal PKSs in the following three paragraphs.

Three fundamental domains are found in all iPKSs in *A. nidulans* like in filamentous fungi in general; β -ketosynthase (KS), acyltransferase (AT), and the acyl carrier protein (ACP). The KS catalyzes the C–C bond formation via decarboxylation reactions through Claisen condensations between thioesters. The ACP domain is responsible for transiently holding the growing acyl chain, hereby allowing the loading of malonyl extender units. The acyl groups are transferred from CoA by AT onto KS and ACP. The iterative use of the three domains results in a non-reduced PK, a β -keto thioester. Additional domains can be present in the PKS allowing the introduction of further chemical complexity.

iPKSs in fungi can, based on their catalytic domains, be classified as non-reducing (NR-PKSs), partially reducing (PR-PKSs), or highly reducing (HR-PKSs) [6]. This is based on their ability to reduce the β -keto carbon. In PR- and HR-PKSs, reduction occurs through the β -ketoreductase (KR) domain that converts the β -ketone to a hydroxyl group. The resulting hydroxyl can go all the way to saturation by elimination of water through the dehydratase (DH) domain followed by hydrogenation by enoyl reductase (ER). In addition, reducing PKSs can also possess a methyltransferase domain (MT) responsible for C-methylation of the growing PK chain, using S-adenosylmethionine (SAM) as a carbon-donor. The degree of modifications and their position in the PK product is always the same for the individual PKSs. However, it is presently unknown how deployment of the various modifying domains is programmed into the PKS enzyme.

NR-PKSs differ in domain architecture from reducing PKS by not having any of the reducing domains and by having an N-terminal starter unit-ACP transacylase (SAT) domain and an internal product template (PT) domain. The SAT domain is responsible for selecting the starter unit to be extended by the enzyme [7], while the PT domain is responsible for folding and cyclization of the non-reduced PK backbone [3,8]. The number of iterations within the PKS and thereby the display of functional groups and the size of the final product is likely determined by the size of the active site cavity in the iPKS [9]. Once the length of the final product has been achieved, the PK chain is released from the PKS, catalyzed by either a thioesterase (TE), a Claisen cyclase (CLC) domain if present, or by accessory enzymes. A more detailed discussion on PKS release mechanisms is reviewed by Du and Lou [10].

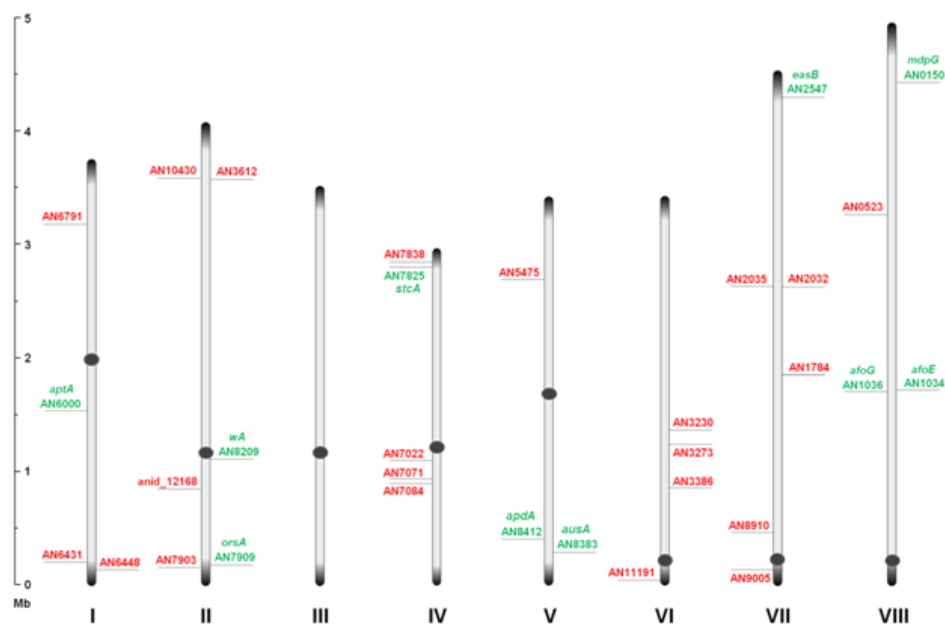
It should be noted that it currently is impossible to reliably predict the product of iPKSs based on their amino acid sequences and domain architecture. This is in part due to the inability to predict the number of iterations performed by the iPKS and in part due the lack of understanding of how deployment of tailoring domains in individual iterations are programmed into the enzyme.

Interestingly, the PKS encoding gene tends to reside in clusters of genes coding for a broad range of enzymatic activities. The compound coming directly from the PKS rarely seem to be the final product in the biosynthesis, but usually undergoes further modifications by tailoring enzymes from small decorations to drastic and large intervention and couplings.

Through inspection of the genome sequence (genome mining), the latest estimate of genes encoding PKSs in *A. nidulans* is 32 open reading frames (ORFs) [11] (Figure 1), indicating that the number of PK containing end products in *A. nidulans* should count at least 32 plus stable intermediates. The compounds detected under a given condition do not necessarily reflect the final outcome of a PK

pathway, since the presence of intermediates and shunt products depends on other downstream enzymes and regulation.

Figure 1. An overview of the relative distribution of the 32 putative polyketide synthases (PKS) open reading frames (ORFs) on the eight chromosomes of *A. nidulans*. Green and red AN numbers represent assigned and unassigned PKS genes, respectively. Dark grey circles and ends symbolize centromeric and telomeric regions, respectively, and should not be considered to scale.



At present, a total of nine PKS genes have been coupled to the polyketome (collection of PKs and their synthesis) in *A. nidulans*, and numerous endeavors are currently attempting to unveil the mechanisms of PK metabolism in this model fungus. The pathways described in this review will follow in chronological order with respect to the discovery from PK to genes. For each of the PK gene clusters that have been linked to products so far we will focus on the PK compounds, their discovery, genetics as well as their biosynthetic pathway: Naphthopyrone, sterigmatocystin, aspyridones, emericellamides, asperthecin, asperfuranone, monodictyphenone (emodin), orsellinic acid, and the austinols.

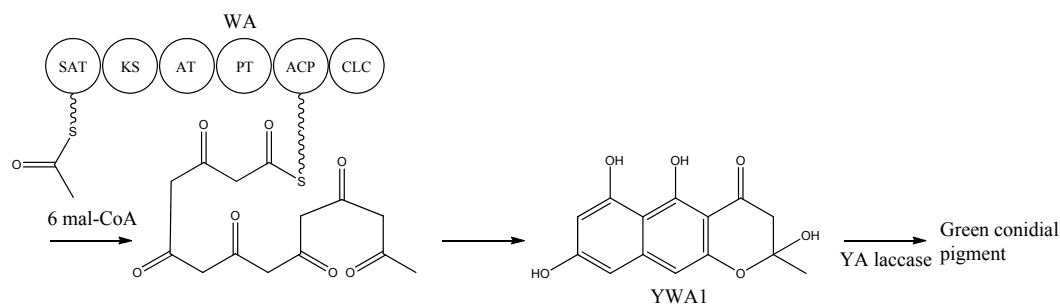
2. Naphthopyrone

Spores from *A. nidulans* are characterized by a dark grey-green macroscopic appearance. This is due to deposition of pigments in the conidial wall as shown by ultrastructure studies using transmission electron microscopy (TEM) [12]. The responsible pigment is based on the PK-naphthopyrone YWA1 and the function of the pigment layer has been shown to include quenching of reactive oxygen species [13] and increased resistance to UV radiation [14]. The work on naphthopyrone synthesis in *A. nidulans* has paved the way for understanding iPKS domain structure.

The study of conidial pigmentation in *A. nidulans* has been extremely valuable for genetic screens. The first pigment mutant recorded in literature was the spontaneous white alba (w_a) strain reported by

Yuill in 1939 [15]. In addition to the white mutant class, yA^- mutants producing yellow conidia were discovered. These available color variants served as easy recognizable markers (green, white, and yellow) that allowed the establishment of fundamental genetic tools in *A. nidulans* [1]. Interestingly, sexual crossing showed that the wA^- mutation masked the effect of the yA^- mutation (epistatic) [1,16]. Clutterbuck and co-workers [16] proposed that WA synthesized the yellow pigment that was observed in yA mutants and that the YA enzyme converted this compound into the green conidial pigment. In 1967, Agnihotri and co-workers [17] found that the wild type strains if grown under copper limiting conditions could mimic the yellow phenotype of the yA^- strain. yA (AN6635) and wA (AN8209) were isolated and mapped to loci, chromosome I and II respectively, by complementation of a cosmid based library in 1989 and 1990, respectively [18,19]. Later, cross-feeding experiments performed by Clutterbuck [16] revealed that the yA phenotype was caused by the lack of a copper dependent extracellular laccase (p-diphenol oxidase). The wA functionality in pigment formation was confirmed by gene-deletion studies [19]. The lack of clustering of the two *A. nidulans* conidial pigment genes also became evident by their different expression patterns and the finding that they are controlled by different regulatory systems [20,21]. The yA gene is expressed in phialide cells and primary sterigmata (metulae) [18], and controlled by BrlA and AbaA [21], while wA is expressed only in phialides [22] and controlled by WetA [20]. Interestingly, none of the genes are expressed in the conidia. Characterization of the WA PKS was accomplished by Northern blotting, which revealed that wA encoded a 7.5 kb large transcript [19], and sequencing of the locus [22]. Re-sequencing of the 3' region in 1998 led to a revised gene model of the PKS with the following domain structure KS-AT-ACP-CLC. This novel CLC domain [23] catalyzed release of the product and cyclization of the second aromatic ring of YWA1 via a Claisen condensation reaction [24].

Heterologous expression of wA in *A. oryzae* resulted in the production of the yellow compound, as observed in yA^- mutants, which was identified to be the heptaketide naphthopyrone named YWA1 [25]. In 2002, wA was used for constructing a collection of chimeric PKSs (cPKSs) by mixing its domains with those of *Colletotrichum lagenarium pks1*, known to produce the tetraketide 1,3,6,8-tetrahydroxynaphthalene (T4HN). One of the resulting cPKSs, SW-B, produced several new compounds including both *tetra*- and *pentaketides* [26]. The results prompted a reanalysis of the two PKSs, which revealed the existence of two previously overlooked conserved domains; an N-terminal and a central domain. These domains were later identified as a SAT and PT domain, respectively, thus providing the full domain structure SAT-KS-AT-PT-ACP-CLC [3,7,8]. With the organization within WA in mind, the biosynthetic pathway can be envisioned as condensations of an acetyl-CoA with six malonyl-CoA units in six successive reactions resulting in the formation of YWA1 [25] (Figure 2).

Figure 2. Biosynthetic pathway for formation green conidial pigment in *A. nidulans*

YWA1 is then believed to be dimerized or polymerized by the YA laccase into the green conidial pigment(s) via phenolic oxidative coupling. However, to date no one has succeeded in characterizing the chemical structure of the green conidial pigmentation in detail. As the final product remains elusive, it is impossible to predict if other tailoring enzymes further modify the YWA1 backbone or reactions occur with other metabolites or cellular components, e.g., the cell wall.

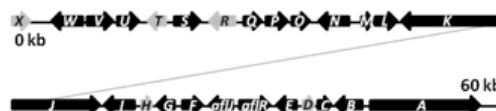
3. Sterigmatocystin

Sterigmatocystin, a PK, was first partially purified from a *Sterigmatocystis* sp. in 1948 by Nekam and Polgar [27]. Hatsuda and co-workers [28,29] successfully isolated sterigmatocystin in 1954 from *A. versicolor*. The correct relative structure was determined in 1962 by Bullock *et al.* [30]. By performing degradative experiments it was shown that the stereochemistry of sterigmatocystin was the same as that of aflatoxin [31], which had the absolute stereochemistry determined in 1967 [32]. The absolute stereochemistry of sterigmatocystin was confirmed via crystallography [33,34].

The aflatoxins are among the most carcinogenic mycotoxins and the research in aflatoxin and sterigmatocystin intensified with the Turkey X disease caused by aflatoxins in the middle of the last century [35]. Aflatoxins are reported to be produced only by a few aspergilli. *A. nidulans* does not produce aflatoxins, as the biosynthesis stops at sterigmatocystin, a late, yet stable precursor of the pathway. Sterigmatocystin is a powerful mycotoxin, though it is estimated to be 150 times less carcinogenic than the most potent aflatoxin, B₁ [36]. Fungi that are able to produce aflatoxins and/or sterigmatocystin are common contaminants of food, feed, and indoor environments and may be mammalian and plant pathogens [37,38]. Due to the high toxicity and prevalence of sterigmatocystin and aflatoxins, they are likely the most extensively studied examples of secondary metabolism in fungi both in terms of biosynthesis and biological function, and there are several excellent and comprehensive reviews for further reading on aflatoxin biosynthesis [39,40]. Studies on the biosynthesis of aflatoxin and sterigmatocystin have been carried out in several fungi (*A. flavus*, *A. nidulans* and *A. parasiticus*) and some of the assigned gene functions in *A. nidulans* are proposed based on gene homology to these two other species.

The biosynthetic cluster of sterigmatocystin in *A. nidulans* was first characterized by Brown and co-workers in 1996 [41]. They identified a 60 kb region in the *A. nidulans* genome responsible for the synthesis of sterigmatocystin. The cluster contains 27 genes named *stcA-X* (Figure 3), reflecting their order of appearance on the chromosome [41].

Figure 3. The sterigmatocystin gene cluster. The black arrows are predicted *stc* ORFs, while light grey arrows are genes with unassigned functions.



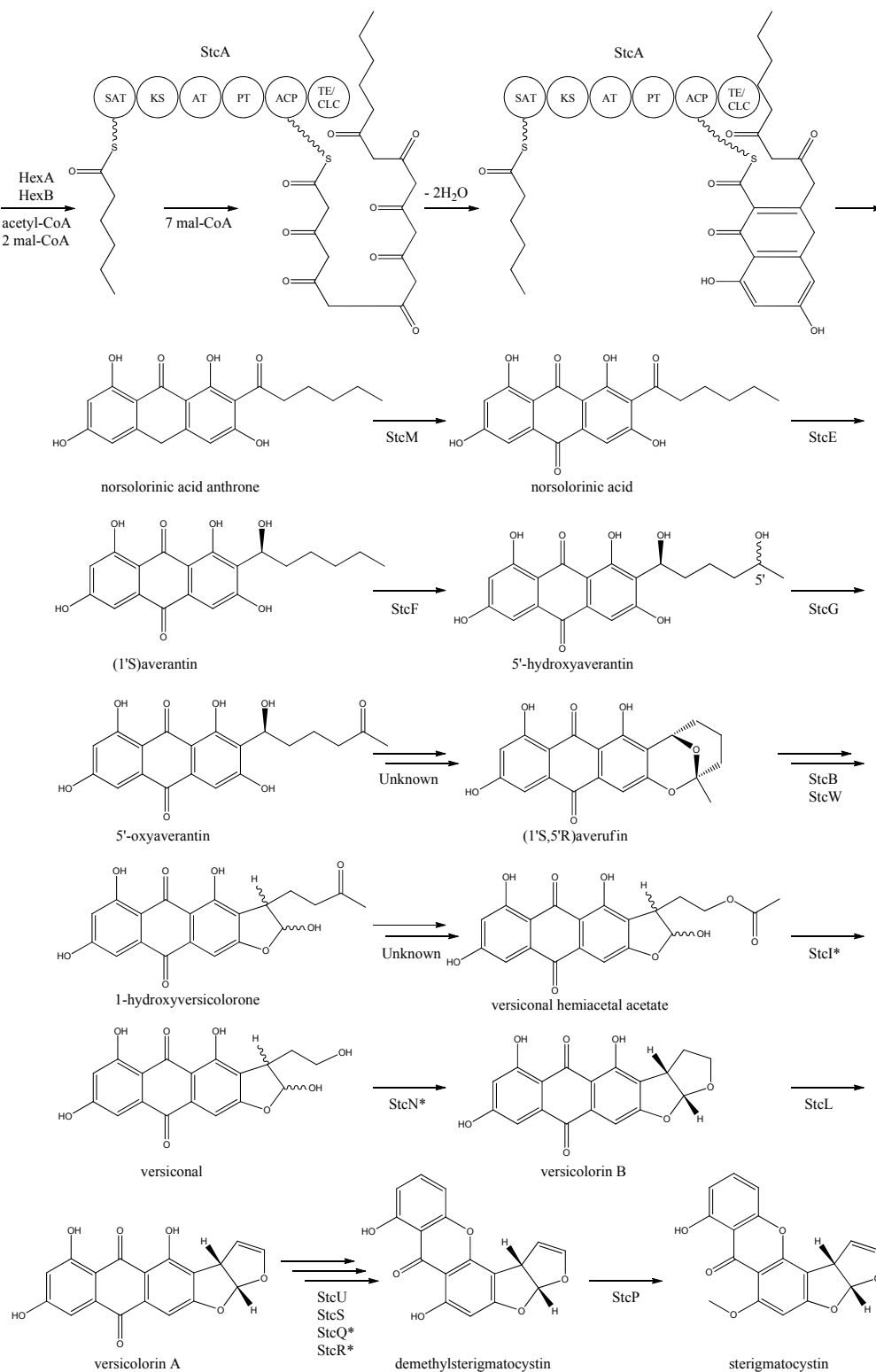
The PKS catalyzing the production of the PK backbone of sterigmatocystin was identified by Yu and co-workers in 1995 [42] and originally named *pksST*, but was later renamed to *stcA* by Brown *et al.* [41] to simplify nomenclature. Besides the PKS, the *stc* gene cluster is predicted to contain two transcription factors (*aflR*, *aflJ*), six monooxygenases (*stcB*, *stcC*, *stcF*, *stcL*, *stcM*, *stcS*, *stcW*), two dehydrogenases (*stcG*, *stcV*), an esterase (*stcI*), an O-methyltransferase (*stcP*), two ketoreductases (*stcE*, *stcU*), a VERB synthase (*stcN*), an oxidase (*stcO*), a monooxygenase/oxidase (*stcQ*), a Baeyer-Villiger oxidase (*stcR*), a fatty acid synthase composed by the two subunits HexA and HexB (encoded by *stcJ* and *stcK*, respectively), and five unassigned genes (*stcD*, *stcH*, *stcR*, *stcT*, *stcX*), which may also be part of the cluster [39–41,43–47].

The *stc* cluster is a relatively large gene cluster, and studying the gene regulation has led to important discoveries. The two TFs were found to be present within the cluster. The AflR is a Zn₂Cys₆ TF that regulates transcription of the *stc* locus in *A. nidulans* [41,48], while AflJ (also named AflS) have been shown to have a role in the regulation of aflatoxin biosynthesis in *A. flavus* and is likely to have a similar function in *A. nidulans* [49]. Interestingly, Bok and Keller [50] discovered a novel regulator (LaeA) of secondary metabolism in *A. nidulans* in a mutant screen for loss of *aflR* expression. Deletion of *laeA* resulted in a significantly decreased production of different classes of SMs like sterigmatocystin and penicillin. LaeA, a putative methyl transferase was moreover acting in a feedback loop with AflR since overexpression of *aflR* downregulates *laeA* expression, and overexpression of *laeA* could not increase production of sterigmatocystin [50]. LaeA was shown to be a part of the conserved Velvet complex, which is important for regulation of fungal development and secondary metabolism [51]. Another hint on chromatin regulated gene expression came from the deletion of a histone deacetylase, *hdaAΔ*, which led to significant increase in the expression of two *stc* cluster genes, *stcU* and *aflR* compared to the reference [52].

Applying this strategy of deleting and overexpressing genes encoding global epigenetic regulators has paved the way for novel discoveries in secondary metabolism. Moreover the alternative of utilizing a chemical epigenetic approach through epigenetic modifier molecules has proven successful in activating gene clusters in *A. niger* [53].

The first step in the biosynthesis of sterigmatocystin (Figure 4) is the production of hexanoate by the FAS units, StcJ and StcK [41]. Watanabe and Townsend [54] showed that the hexanoyl-CoA is not an intermediate freed from the complex, indicating that hexanoate is transferred directly to the SAT domain of the PKS. The PK backbone is assembled by StcA by condensation of the starter unit, hexanoyl-CoA and seven malonyl-CoA extender units followed by cyclization and release of norsolorinic acid anthrone [42]. The oxidation of norsolorinic acid anthrone to norsolorinic acid may be catalyzed by *stcM*, a monooxygenase ortholog to *hypC* that converts norsolorinic acid anthrone to norsolorinic acid in *A. parasiticus* [43]. Norsolorinic acid is the first stable intermediate in the biosynthesis of sterigmatocystin and is converted into averantin by StcE, reducing the hexanoate ketone to an alcohol [41,55,56].

Figure 4. Proposed biosynthesis of sterigmatocystin. StcA contains starter unit-ACP transacylase (SAT), β -ketosynthase (KS), acyltransferase (AT), product template (PT), acyl carrier protein (ACP) and thioesterase/claisen cyclase (TE/CLC) domains. *Indicates a proposed, but not confirmed, enzyme. Multiple arrows indicate that the number of enzymatic steps is unknown.



Yabe *et al.* [55] showed that 5'-hydroxyaverantin is a step towards aflatoxin in *A. parasiticus* and a later study showed that the oxidation of averantin into 5'-hydroxyaverantin is catalyzed by StcF [57]. The conversion of 5'-hydroxyaverantin to 5'-oxyaverantin is likely catalyzed by StcG [39,44]. In a study in *A. parasiticus* by Yabe and co-workers it was shown that both (1'S, 5'S)- and (1'S, 5'R)-hydroxyaverantin are formed in the conversion of averantin to 5'-oxyaverantin [58]. Disruption of *aflH* in *A. parasiticus* resulted in the accumulation of 5'-hydroxyaverantin, however, small amounts of *O*-methylsterigmatocystin present suggested that other enzymes may be involved in the reaction [44,59]. The gene(s) responsible for the conversion of 5'-oxyaverantin to averufin have not been identified [39,60].

Individual disruption of StcB and StcW resulted in elimination of sterigmatocystin and accumulation of averufin, indicating that both enzymes catalyze the conversion of averufin to 1-hydroxyversicolorone [57]. It was not possible for the authors to determine why two monooxygenases were required for this reaction step [57]. No gene products have been identified as being responsible for the conversion of 1-hydroxyversicolorone to versiconal hemiacetal acetate. StcI is thought to catalyze the reaction from 1-hydroxyversicolorone to versiconal based on studies of the ortholog AflJ in *A. parasiticus*, though other genes capable of this reaction may be present in *A. nidulans* [44,61].

Deletion of *stcN* did not result in the production of sterigmatocystin or other intermediates [44]. However, StcN show homology to AflK and the versicolorin B synthase, Vbs, in *A. parasiticus*, indicating that the biosynthetic step from versiconal to versicolorin B may be catalyzed by StcN [39,44,62]. StcL was shown by Kelkar *et al.* [63] to catalyze the conversion of versicolorin B to versicolorin A. Inactivation of *stcL* resulted in accumulation of dihydrosterigmatocystin, leading to a branching of the sterigmatocystin biosynthesis as seen in the aflatoxin biosynthesis. Addition of versicolorin A to the mutant gave production of sterigmatocystin, and that indicated that this enzyme functions before versicolorin A [63].

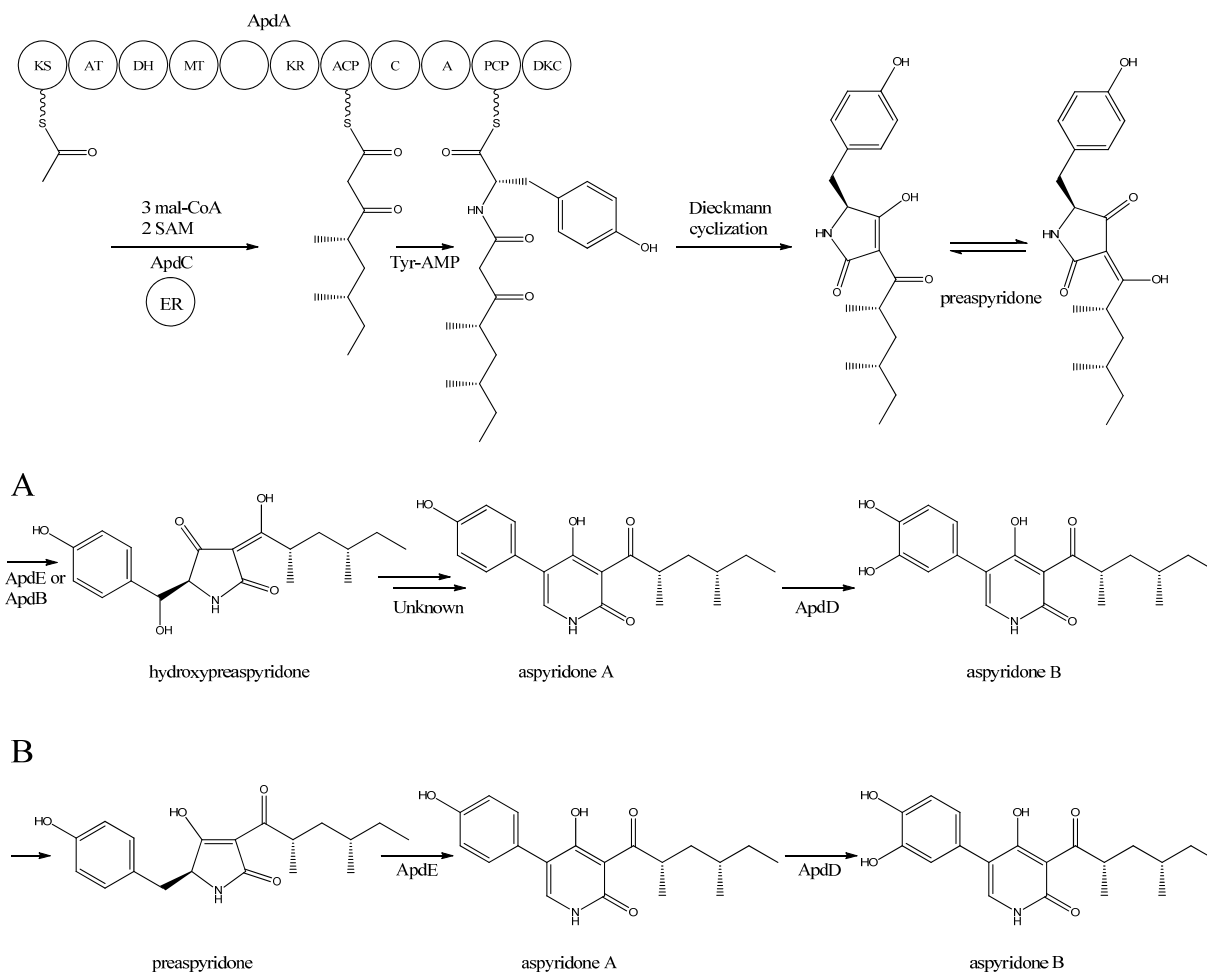
Keller and co-workers [64,65] showed that StcU and StcS are involved in the conversion of versicolorin A to demethylsterigmatocystin. Individual disruption of *stcU* and *stcS* led to the accumulation of versicolorin A and eliminated production of sterigmatocystin in *A. nidulans* [64,65]. Henry and Townsend [45] studied the same step in the aflatoxin biosynthesis in *A. parasiticus* and proposed an oxidation-reduction-oxidation mechanism, involving at least a ketoreductase AflM and a monooxygenase AflN, orthologs to *stcU* and *stcS*, respectively. Ehrlich *et al.* [46] and Cary *et al.* [47] identified two enzymes, AflY, a Baeyer-Villiger oxidase and AflX, an oxidoreductase, to be involved in the conversion of versicolorin A to demethylsterigmatocystin in *A. parasiticus* and *A. flavus*, respectively. *aflX* and *aflY* are homologous to *stcQ* and *stcR*, which suggests that these genes might be involved in the biosynthetic step from versicolorin A to demethylsterigmatocystin [46,47].

The final step in the biosynthesis of sterigmatocystin is the methylation of demethylsterigmatocystin catalyzed by StcP [66]. The conversion of sterigmatocystin to aflatoxin involves two additional biosynthetic steps; an *O*-methylation of sterigmatocystin by *aflP* followed by involvement of *aflQ* to produce aflatoxin [39]. Slot and Rokas have recently showed that the sterigmatocystin gene cluster in *Podospira anserina* was horizontally transferred from *Aspergillus*, which shows that transfer of large metabolite clusters between fungi are possible [67].

4. Aspyridone

Aspyridone is a PK-NRP hybrid and a fascinating example on how PK-NRP compounds in *A. nidulans* can be assembled from the activity of a single fusion enzyme. The aspyridones have shown to display moderate cytotoxicity [68]. The responsible gene cluster was discovered by Bergmann and co-workers [68] using a genome mining approach. Using the *Aspergillus* genome sequence, they identified a SM gene cluster, which contained a putative TF (AN8414/*apdR*) that the authors hypothesized could trigger activation of the genes in the cluster. Accordingly, the authors overexpressed the TF under the control of an inducible *alcA* promoter by integrating it randomly in the genome. In agreement with the hypothesis, it was demonstrated by Northern blot analysis that six of the nearest neighbor genes were up-regulated in this strain on inductive medium, and that the aspyridones and two intermediates or shunt products could also be detected. Prediction of the catalytic potential for the six upregulated genes (*apdA*, *apdB*, *apdC*, *apdD*, *apdE*, and *apdG*) combined with the structure of the accumulating compounds allowed the authors to propose a model for the biosynthesis of aspyridones including an assignment of the involved enzymes (Figure 5).

Figure 5. Proposed biosynthesis of aspyridone A and B. **(a)** Based on proposed biosynthesis by [68,72,73]; **(b)** Based on results by Halo *et al.* [78] Multiple arrows indicate that the number of enzymatic steps is unknown.



apdA was deleted by Chiang *et al.* [69] and confirmed to be involved in aspyridone biosynthesis as reported by Bergmann *et al.* [68].

The structure of aspyridone A and B suggested that their synthesis involved both PKS and NRPS activity. Indeed, analysis of the AN8412 structure revealed domains characteristic for a HR-PKS as well as NRPS in one ORF spanning more than 11 kb. This is a special subclass of reducing PKSs, where the PKS has been directly fused with a single NRPS module at the C'-terminal end. This architecture allows for the incorporation of amino acids or carboxylic acids into the carboxylic end of the growing PK chain. Only one of these fusion enzymes has been found in *A. nidulans*, but has been reported in other fungi [70,71]. Since AN8412 is the first enzyme to act in the pathway, the gene was named *apdA*. ApdA catalyzes the assembly of the PK-amino acid backbone of the aspyridones by three Claisen condensations of malonyl-CoA, and KR-DH-ER-MT carries out full reduction of the β -keto and the methylations, which are required. However, as ApdA lacks a functional ER domain, the ER activity is most likely provided by ApdC, a homolog to an enoyl reductase (LovC) from the lovastatin biosynthetic gene cluster [68]. The resulting triketide is transferred to the NRPS module, where it is linked to tyrosine [68]. Bergmann and co-workers listed the domains through protein homology in the NRPS as condensation (C), adenylation (A), peptidyl carrier protein (PCP) and reductase domain (RED). The release of the PKS-NRPS hybrid product was proposed to be a NADPH-dependent reductive release followed by an intramolecular Knoevenagel condensation and enzymatic oxidation [68]. Biochemical studies of the role of ApdA and ApdC in the biosynthetic pathway of the aspyridones have been performed by Liu *et al.* [72] and Xu *et al.* [73]. Liu and co-workers [72] defined the NRPS module as C-A-T-R with the latter two being thiolation and reductase, which is an alternative to the more frequent C-A-T-TE found in these modules. However, this reductase domain (R*) in the NRPS module of ApdA is not the standard SDR superfamily dehydrogenase since tyrosine in the Ser-Tyr-Lys catalytic triad is mutated suggesting a redox-independent condensation reaction and the release of a tautomer of preaspyridone from ApdA by a Dieckmann cyclization, which was first shown by Halo and co-workers [74]. This result has been confirmed by Xu *et al.* who expressed the *apdA* and *apdC* genes in *Saccharomyces cerevisiae* and *Escherichia coli*, respectively [73]. The purified enzymes (ApdA and ApdC) were incubated in the presence of cofactors and building blocks and the predominant product was preaspyridone [73].

The formation of preaspyridone into aspyridone A and B was proposed by Bergmann *et al.* [68], (and outlined in Figure 5a) using the predicted functions of the remaining genes of the *apd* gene cluster. The proposed biosynthesis involved ApdB and ApdE which shows similarity to cytochrome P450 oxygenases and cytochrome P450 alkane hydroxylases, respectively, and were believed to catalyze the formation of hydroxypreaspyridone [68]. Based on the study of pyridone rearrangement in metabolites related to aspyridone it was suggested that ApdE or ApdB were involved in the pyridone rearrangement [68,75–77]. Moreover, aspyridone A was hypothesized to be converted into aspyridone B by ApdD, a putative FAD-dependent monooxygenase, which is related to other ring hydroxylases [68]. Aspyridone has a similar structure to other pyridines isolated from fungi, e.g., tenellin whose biosynthetic gene cluster also has been identified [75,76]. The proposed biosynthesis of aspyridone was, as described above, based on predicted gene functions and not isolated intermediates. However, a study on the biosynthesis of the related metabolite tenellin by Halo and co-workers [78] showed that the suggested biosynthesis may be incorrect and an alternative biosynthesis was suggested (as shown

in Figure 5b). In this biosynthesis preaspyridone is not converted into hydroxyaspyridone but ring expanded by ApdE to aspyridone A similar to the biosynthesis of tenellin [78]. Halo *et al.* also showed that the hydroxylated metabolite of pretenellin is a shunt metabolite as it could not be converted into tenellin.

5. Emericellamides

The emericellamides are other examples of hybrid compounds that are formed between PKs and NRPs. In this case the biosynthesis requires a PKS and a NRPS rather than a fusion PKS-NRPS as used in the production of the aspyridones. Emericellamides are cyclic depsipeptides and a total of five variants, A, C–F, of these metabolites have been found in *A. nidulans*. Initially, emericellamide A and B were isolated and described from an unidentified marine-derived *Emericella* strain in a screen due to their antibacterial activity against methicillin-resistant *Staphylococcus aureus* [79].

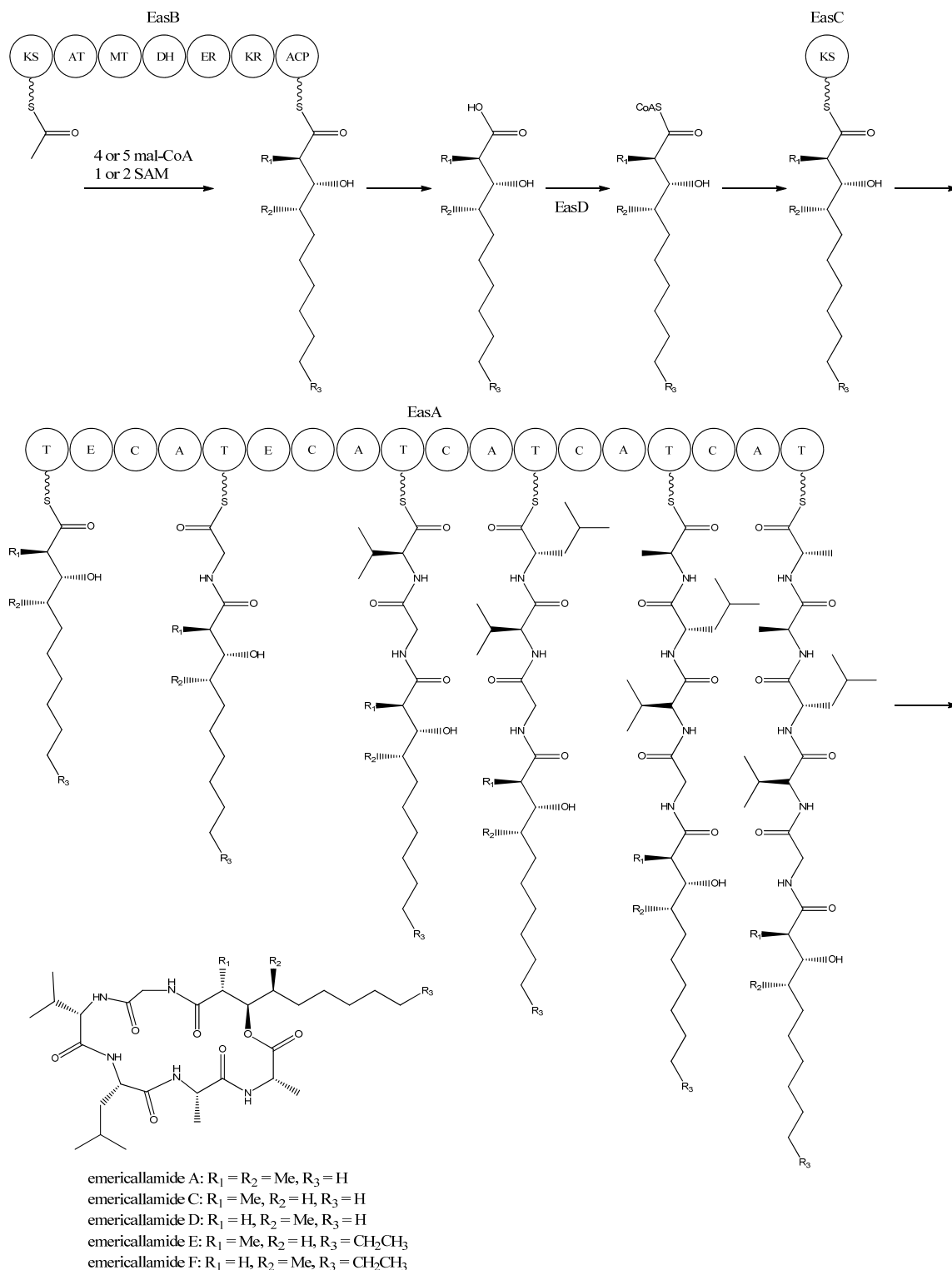
In order to discover novel natural products, Chiang and co-workers [69] searched the genome sequence of *A. nidulans* for NRPS gene candidates. Subsequently, six of these genes were randomly chosen and deleted by gene targeting. One of the resulting mutants, AN2545Δ, showed a metabolite profile where emericellamide A was missing. Furthermore, HPLC profiles and dereplication using mass spectrometry and database searches revealed four additional compounds, which disappeared in the mutant metabolite profile. Since these compounds had not previously been described in *A. nidulans*, they were purified and their structures solved by NMR analysis revealing that they were novel analogues of emericellamide A and B, thus named emericellamide C–F [69].

To investigate whether AN2545, now called *easA*, defines a gene cluster encoding all necessary enzymatic activities in the emericellamide biosynthetic pathway, the genes from AN2542 to AN10325, a total of ten, were deleted [69]. Most of these gene deletions did not affect emericellamide production as judged by LC-MS analysis, demonstrating that they do not participate in the biosynthesis. However, the emericellamides were absent in four of the deletion strains, now named *easA-easD*, indicating that these genes are involved in the pathway.

Bioinformatic analysis of the three additional genes suggested that they all encode activities that are relevant for emericellamide biosynthesis. Specifically, *easB* (AN2547), a PKS, *easC* (AN2548), an acyl transferase, and *easD* (AN2549), an acyl-CoA ligase. Based on these putative activities, the authors proposed a biosynthetic pathway for emericellamide production (Figure 6). In this model, the biosynthesis is initiated by EasB, a HR-PKS composed of the domains KS-AT-MT-DH-ER-KR-ACP. Since the PK component of the different emericellamide variants differ with respect to chain length and methylation pattern, it indicates that iterativity of this PKS is flexible [69].

Next, the PK carboxylic acid is converted to a CoA thioester by the acyl-CoA ligase, EasD, loaded onto the acyltransferase EasC, and then transferred to the thiolation (T) domain of EasA. This NRPS is a multi-modular enzymatic assembly containing 18 domains grouped into five modules. Among those, the authors propose that the first T domain is responsible for accepting the incoming PK from EasC (Figure 6). Moreover, the remaining domains fit well with the fact that five amino acids are incorporated into emericellamides. The authors note that this NRPS does not contain a TE domain at the end of module 5, indicating that this enzymatic activity is not necessary for cyclization of the emericellamides [69].

Figure 6. Proposed biosynthesis of the emericallamides. The order of the methyltransferase (MT) and dehydratase (DH) domain as suggested by Chiang *et al.* [69], however a BLASTp analysis suggests a swapping of the MT and DH domains. The NRPS, EasA, contains 18 (T, E (epimerization), C, A) domains grouped into five modules.



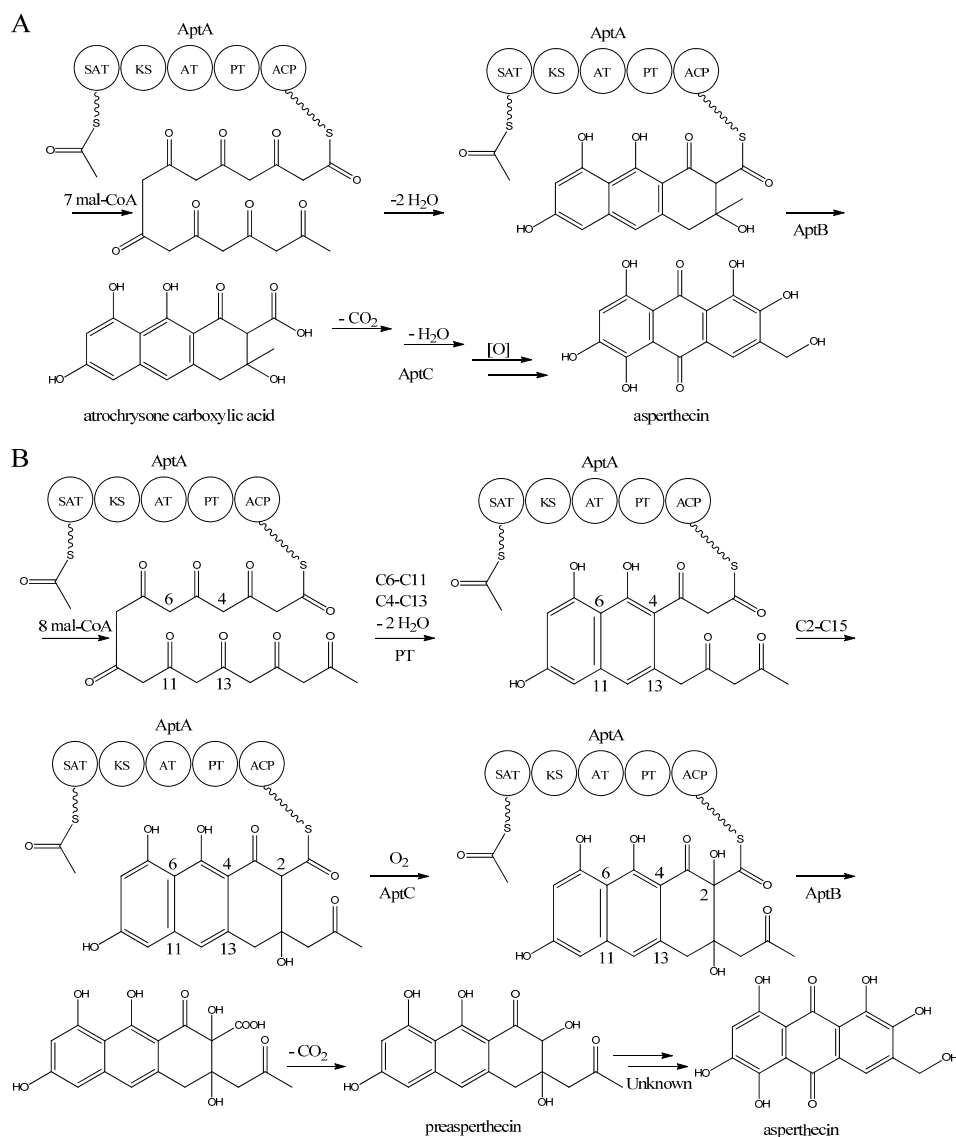
6. Asperthecin

Asperthecin is a PK compound that was first isolated from *A. quadrilineatus* by Howard and Raistrick in 1955 [80], however, the structure was not fully determined until six years later [81]. Initially, various chemical derivatizations and spectroscopic data determined a partial structure of asperthecin [80]. Neelakantan and co-workers [82] reduced the number of possible structures to two, and further derivatizations of asperthecin by Birkinshaw and Gourlay resulted in the final structure [81]. *A. quadrilineatus* is a member of the *A. nidulans* group, therefore Howard and Raistrick [80] extended the search of asperthecin to additional members of the *A. nidulans* group. No other was as rich in asperthecin production as *A. quadrilineatus*, yet small amounts of crystallized asperthecin could be obtained from cultures of *A. nidulans* and *A. rugulosus* indicating that production of asperthecin was possible in other aspergilli [80].

About 50 years later, Szewczyk and co-workers [83] used a molecular genetics approach to find the gene cluster responsible for the production of asperthecin in *A. nidulans*. Since many aspects of the regulation in the polyketome were unknown, the authors speculated whether sumoylation had an effect. SUMO is a small ubiquitin-like protein which is post-translationally added to proteins in the cell, as it plays a role in regulating transcription. *A. nidulans* contains one SUMO encoding gene, *sumO* [84], deletion of which led to a decrease in the production of SMs such as austinol, dehydroaustinol, and sterigmatocystin, and an increase in the production of a metabolite identified to be asperthecin, whereas the production of emericellamides were not affected [83]. Due to the aromatic structure of asperthecin, Szewczyk *et al.* [83] studied the domain prediction in 27 putative PKS-protein sequences using the *A. nidulans* genome sequence and available tools, in order to identify potential producers of non-reduced PKs. Ten NR-PKSs were identified and a deletion series of all NR-PKS genes was performed in the *sumOΔ* background [83]. While nine of the PKS-deletion strains still produced asperthecin, the AN6000 (*aptA*) PKS-deletion strain failed to synthesize asperthecin [83]. With the notion that most end compounds in PK biosynthesis are made by a clustered gene collective, six candidate genes surrounding *aptA* were picked in an attempt to identify the *apt* biosynthetic cluster. Two of these genes, *aptB* (AN6001) and *aptC* (AN6002), were found to be required for asperthecin production [83]. One strain (AN5999Δ) had a significantly lower production of asperthecin compared to the reference strain, but asperthecin was still present in the metabolite extracts, and as the strain showed poor growth, it was not included in the *apt* gene cluster.

Interestingly, AptA was shown to have SAT-KS-AT-PT-ACP domains, but lack a TE/CLC domain [83,85–88]. Independent groups have used this case as a model system to study the mode of PK release, and two alternating mechanisms for the biosynthesis of asperthecin are shown in Figure 7 [86,87]. The first model suggests the formation of the PK backbone by condensations of one acetyl-CoA and seven malonyl-CoA units [86], with the β -lactamase AptB releasing the octaketide from AptA [83]. This assumption was based on a study by Awakawa and co-workers [85] in *A. terreus*, where there was a release of atrochrysone carboxylic acid from the atrochrysone carboxylic acid synthase (ACAS) lacking a TE/CLC domain, in the presence of the atrochrysone carboxylic ACP thioesterase (ACTE), a member of the β -lactamase superfamily [85]. The unstable atrochrysone carboxylic acid then undergoes a series of reactions; decarboxylation, dehydration, and various oxidations where the monooxygenase AptC is believed to be involved, and in the end yielding asperthecin [86].

Figure 7. Proposed biosynthesis of asperthecin (a) As suggested by Chiang *et al.* [86]; (b) Based on proposed biosynthesis by Li *et al.* [87]. AptA contains SAT, KS, AT, PT and ACP domains. Multiple arrows indicate that the number of enzymatic steps is unknown.



In another approach, Li *et al.* [87] introduced *aptA*, *aptB*, and *aptC* into *S. cerevisiae*, which resulted in the production of a nonaketide (here called preasperthecin), and not the octaketide as proposed by Chiang *et al.* [86]. Expressing *aptA* and *aptB* without *aptC* resulted in a product identical to preasperthecin except for the lack of C2-OH, confirming that AptC is responsible for this step [87]. Expression of *aptA* and *aptC* alone did not lead to the production of preasperthecin or any other traceable compounds, confirming that AptB is needed for release of the PK from AptA [87]. These results were confirmed by an *in vitro* assay after expressing *aptA*, *aptB* and *aptC* in *E. coli* [87]. Further insight into AptA functionality came from expressing the AptA-PT domain in *E. coli*, and combining it with the *Gibberella fujikuroi* PKS4, which can produce nonaketide products *in vitro* [88]. The experiment revealed that AptA-PT can catalyze C6-C11 cyclization, and most likely also the C4-C13 cyclization. Further, a spontaneous C2-C15 cyclization was followed by a C1-C17

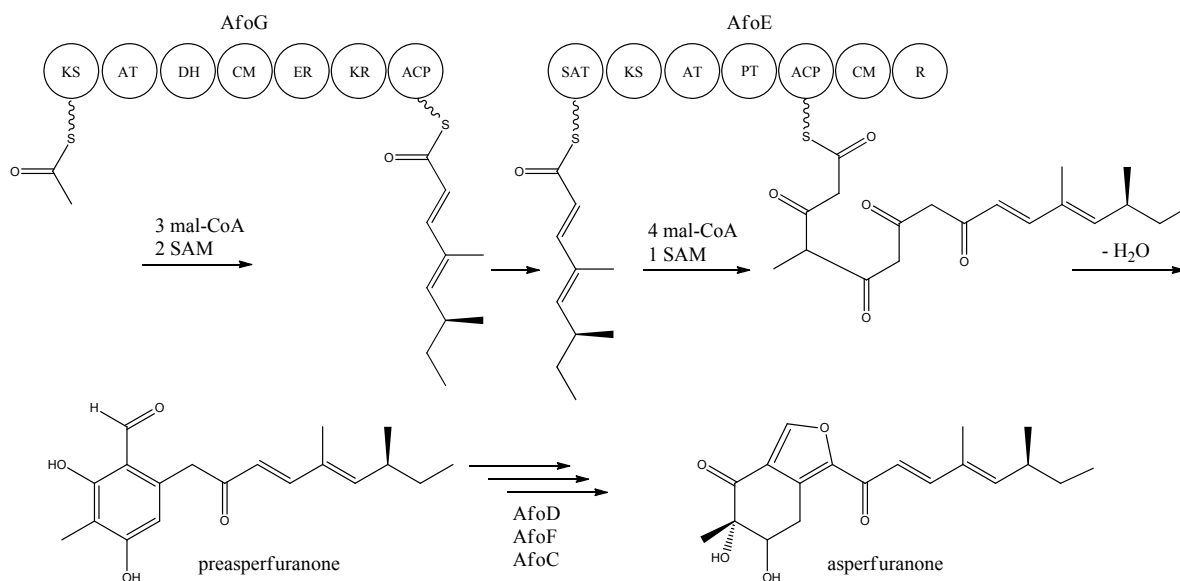
esterification [88]. The *apt* gene cluster appears likely to consist of additional genes, which are responsible for the conversion of preasperthecin into asperthecin.

7. Asperfuranone

Asperfuranone is an example of a novel PK metabolite discovered through genetic mining in *A. nidulans*, as this compound had not previously been reported in the literature before Chiang and co-workers [89] in 2009. Asperfuranone was later shown to possess bioactive properties as it inhibited proliferation of human non-small A549 cancer cells [90]. Investigating the loci containing putative PKS gene clusters, Chiang and co-workers [89] noticed that a NR-PKS gene (AN1034, *afoE*) and a HR-PKS gene (AN1036, *afoG*) were located close to each other on chromosome VIII. Since no products had ever been detected from activity of this locus, and due to the rare constellation of two neighbor PKSs, the authors speculated whether a novel metabolite could be revealed. A putative transcriptional activator (AN1029, *afoA*) was found near the PKS and the authors replaced the upstream sequence of *afoA*, estimated to be the native promoter, with the inducible *alcA* promoter [89]. This indeed turned on the expression of the cluster, since asperfuranone and a precursor metabolite were detected. The structure of asperfuranone was determined based on one- and two-dimensional NMR experiments and the absolute configuration by a modified Mosher's method, whereas the precursor preasperfuranone had already been determined in the literature [89,91]. With these two compounds being identified, a gene-deletion strategy was performed to map the other genes assigned to the *afo* gene cluster, which involved twelve surrounding genes including the two PKSs [89]. Four deletion strains *afoDA*, *afoEA*, *afoFA*, and *afoGA* fully eliminated asperfuranone production whereas *afoBA* and *afoCA* strongly reduced production of asperfuranone [82]. The deletions confirmed that both *afoE* and *afoG* were responsible for the production of asperfuranone, and that the deletion of *afoD*, encoding a putative hydroxylase, resulted in the production of preasperfuranone [89]. Deletion of *afoB* reduced the production of asperfuranone and due to a high homology to efflux pumps, Chiang and co-workers [89] suggested that it was responsible for the transport of asperfuranone out of the cell.

With the gene cluster and predicted functionalities of the gene products defined, a biosynthetic pathway of asperfuranone was proposed (Figure 8) [89]. The assembly of the primary reduced tetraketide is synthesized by AfoG from one acetyl-CoA, three malonyl-CoA, and two SAM. The tetraketide is transferred to the SAT domain of AfoE and extended with four malonyl-CoA and one SAM [89]. The octaketide is released from AfoE after aldol condensation and reductive release from a C-terminal reductase (R) domain, which resembles a reductive release mechanism to generate the aldehydes described by Bailey *et al.* [92], forming the aldehyde preasperfuranone [89]. The biosynthetic steps from preasperfuranone to asperfuranone are uncharacterized and the suggestions are not based on identified metabolites [89]. Accumulation of preasperfuranone in the *afoDA* suggested AfoD to be the next enzyme in the biosynthesis of asperfuranone. The deletions of *afoF*, encoding a putative FAD/FMN-dependent oxygenase and *afoC*, initially believed to code for a homologue to citrinin biosynthesis oxidoreductase, did not reveal any intermediates, the order of reactions and the exact enzymatic functions for AfoF and AfoC have not been determined. In *afoCA*, the production of asperfuranone was not fully eliminated, which Chiang and co-workers [89] suggested could be due to other enzymes catalyzing the reaction, however less efficiently.

Figure 8. Proposed biosynthesis of asperfuranone. The highly reducing (HR)-PKS, AfoG, contains KS, AT, DH, CM, ER, KR, and ACP domains whereas the NR-PKS, AfoE, contains SAT, KS, AT, PT, ACP, CM, and R domains. The only intermediate isolated in the biosynthesis is preasperfuranone [89]. Multiple arrows indicate that the number of enzymatic steps and reaction order is unknown.



Other puzzling discoveries have been made in connection to the asperfuranone production. When trying to activate a cryptic NRPS gene cluster containing two NRPSs, *inpA* (AN3495) and *inpB* (AN3496), by overexpression of a regulatory gene, *scpR* (AN3492), Bergmann *et al.* [93] also activated asperfuranone. This is an interesting example of a regulatory gene located on chromosome II that activates the *afo* cluster located on chromosome VIII [93]. Lui *et al.* [94] have attempted to engineer the production of a new metabolite by swapping the SAT domain of AfoE with the StcA-SAT. This led to the production of a new metabolite though having the same length as the native AfoE product, asperfuranone [94].

8. Monodictyphenone/Emodin

The PK monodictyphenone was first reported in *A. nidulans* in 2005 [95] and the genes behind the production of monodictyphenone were mapped four years later [96]. This discovery not only enabled the establishment of a biosynthesis model for monodictyphenone in *A. nidulans*, it has subsequently revealed that more than ten different stable products among different classes of related polyketides can be linked to monodictyphenone biosynthesis [11,96–98]. These metabolites count monodictyphenone, emodin and the emodin derivatives 2-hydroxyemodin, 2-aminoemodin, ω -hydroxy emodin, and emodic acid. Moreover, the arugosins and prenyl-xanthenes are also coupled to the pathway [11,98]. The compound emodin has been studied for more than a century [99], and is an anthraquinone found in a wide array of both plants and fungi [100,101]. Emodin and several derivatives (e.g. emodic acid) have been shown to possess anti-bacterial and cancer preventive properties [102–106].

The presence of the SM clusters in silent areas of chromosomes, e.g. near telomeres and centromeres, suggests that chromatin remodeling factors can influence the expression of genes responsible for SMs. As rationalized by Bok and co-workers [96], removal of histone-tail methylation could open heterochromatic regions for transcription. The authors deleted an ortholog, *cclA*, to the yeast *BRE2* gene, encoding an enzyme partner of the COMPASS transcriptional regulator complex conserved in eukaryotes, which rendered *A. nidulans* defective in di- and trimethylation of lysine 4 of the histone 3 tails (H3K4). The *cclA* deletion was established in a mutant strain, *stcJΔ*, to avoid interference of high amounts of sterigmatocystin in purification of other metabolites. The effect was striking as the loss of CclA in HPLC analysis showed an altered chemical landscape compared to the *stcJΔ* reference [96].

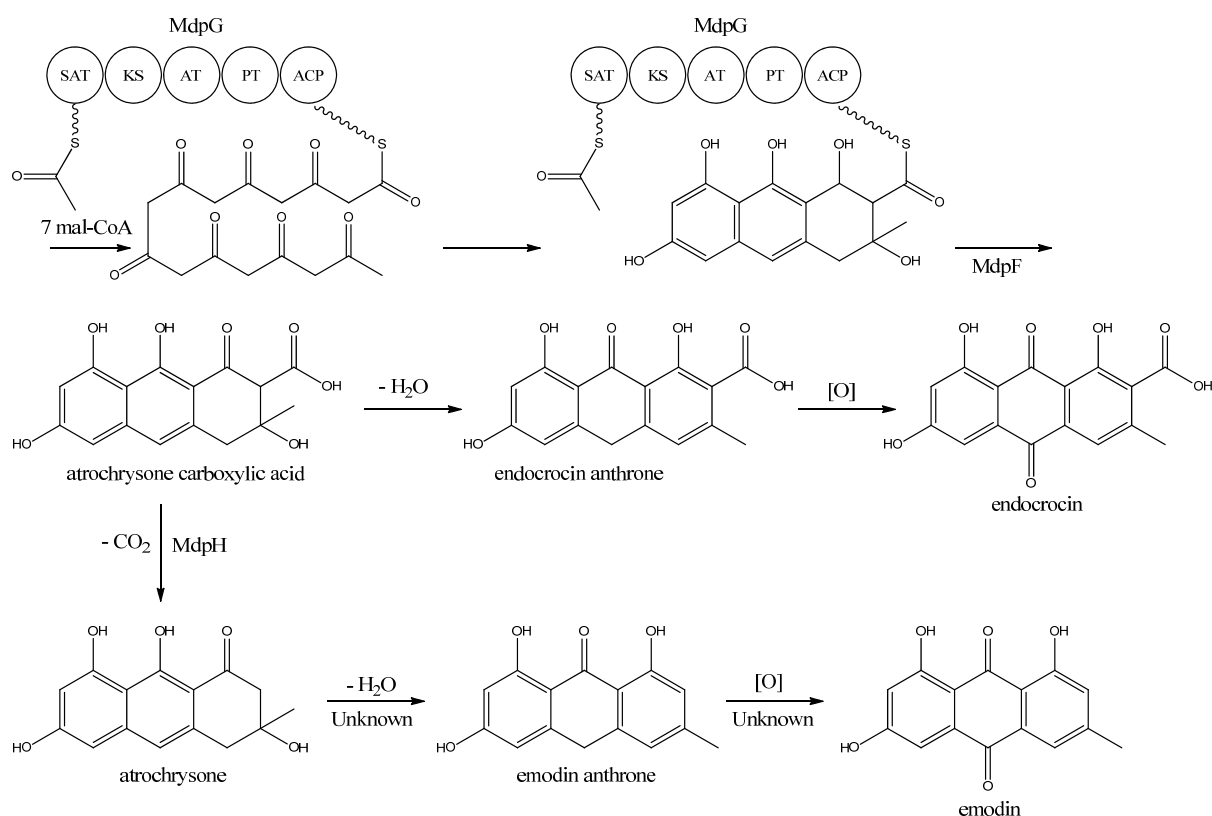
As the compounds appearing were UV-active suggesting high conjugation likely due to aromaticity, the ten NR-PKSs investigated in the asperthecin and asperfuranone studies were individually deleted in the *cclAΔ stcJΔ* double deletion background. This screen revealed six products that all were linked to one PKS, AN0150 (*mdpG*) [96]. Delineation of the cluster was achieved by inspecting the genome sequence for possible cluster candidates followed by Northern blotting for gene-expression analysis in the *cclAΔ stcJΔ*, where the products were detected. The cluster was found to span 12 putative ORFs (AN10021-AN10023 (*mdpA-L*)) [97] of which two genes AN0147 (*mdpD*) and AN10035 (*mdpI*), did not show altered expression from the reference [96,97]. The *mdp*-cluster candidates were also deleted in the *cclAΔ stcJΔ* mutant strain to confirm the expression analysis data and to draw the borders of the cluster [97]. The authors suggest that two transcriptional activators are present within the cluster; MdpE as a main activator (homologue to AfIR) and that MdpA is a co-activator. The *mdp* locus is located near the telomere of chromosome VIII, and activation of the genes in the *cclAΔ* strain supports the hypothesis of epigenetic regulation in these areas through chromatin remodeling [96].

Two groups cultivated *A. nidulans* on complex growth media, which revealed six additional metabolites. First Sanchez *et al.* [98] discovered that emericillin, variecoxanthone A, shamixanthone, and epi-shamixanthone were also products of the *mdp* cluster, and subsequently Nielsen and co-workers [11] added arugosin A and H to the pathway. Since these PKs are prenylated, a BLAST search of the *A. nidulans* genome sequence was performed and pathway-candidate genes were deleted. Two prenyl transferases encoded by *xptA* and *xptB*, and one neighbor GMC oxidoreductase encoded by *xptC*, were found to be involved in the pathway, though they were located on other chromosomes than the *mdp* cluster (for cluster overview see Sanchez *et al.* [98]). This is an intriguing example of SM-cluster members located on more than one chromosome, however, prenyl transferases are known to have broad substrate specificity, and it is currently not known whether they are involved in other processes than prenyl-xanthone formation [98].

MdpG synthesizes the main PK backbone. Since MdpG lacks a CLC/TE domain, MdpF, a putative zinc dependent hydrolase, is believed to catalyze the release of the PK from MdpG [97]. The mechanism is believed to follow the case of ACAS and ACTE as introduced previously in the asperthecin section. Awakawa and co-workers [85] demonstrated that the direct product of the ACAS/ACTE is not emodin anthrone as proposed earlier [107,108], but more likely atrochrysone carboxylic acid (Figure 9). Atrochrysone carboxylic acid was not observed *in vitro*, instead the decarboxylated product atrochrysone was the major product in the assays and therefore proposed to be an intermediate to emodin, as suggested by Couch and Gautier [109]. Conversion of atrochrysone to

emodin requires dehydration (forming emodin anthrone) and a final oxidation (Figure 9), however, Awakawa and co-workers [85] observed small amounts of both emodin anthrone and emodin *in vitro* showing that these reactions may occur non-enzymatically. Based on these observations, Chiang and co-workers [97] proposed that *mdpH* encodes a decarboxylase, catalyzing the conversion of atrochrysone carboxylic acid to atrochrysone. The deletion of *mdpH* resulted in accumulation of a shunt product endocrocin produced via endocrocin anthrone. Enzymes responsible for dehydration of atrochrysone or modification of emodin into the observed derivatives have not yet been identified.

Figure 9. Proposed biosynthesis of emodin. MdpG contains SAT, KS, AT, PT, and ACP domains.

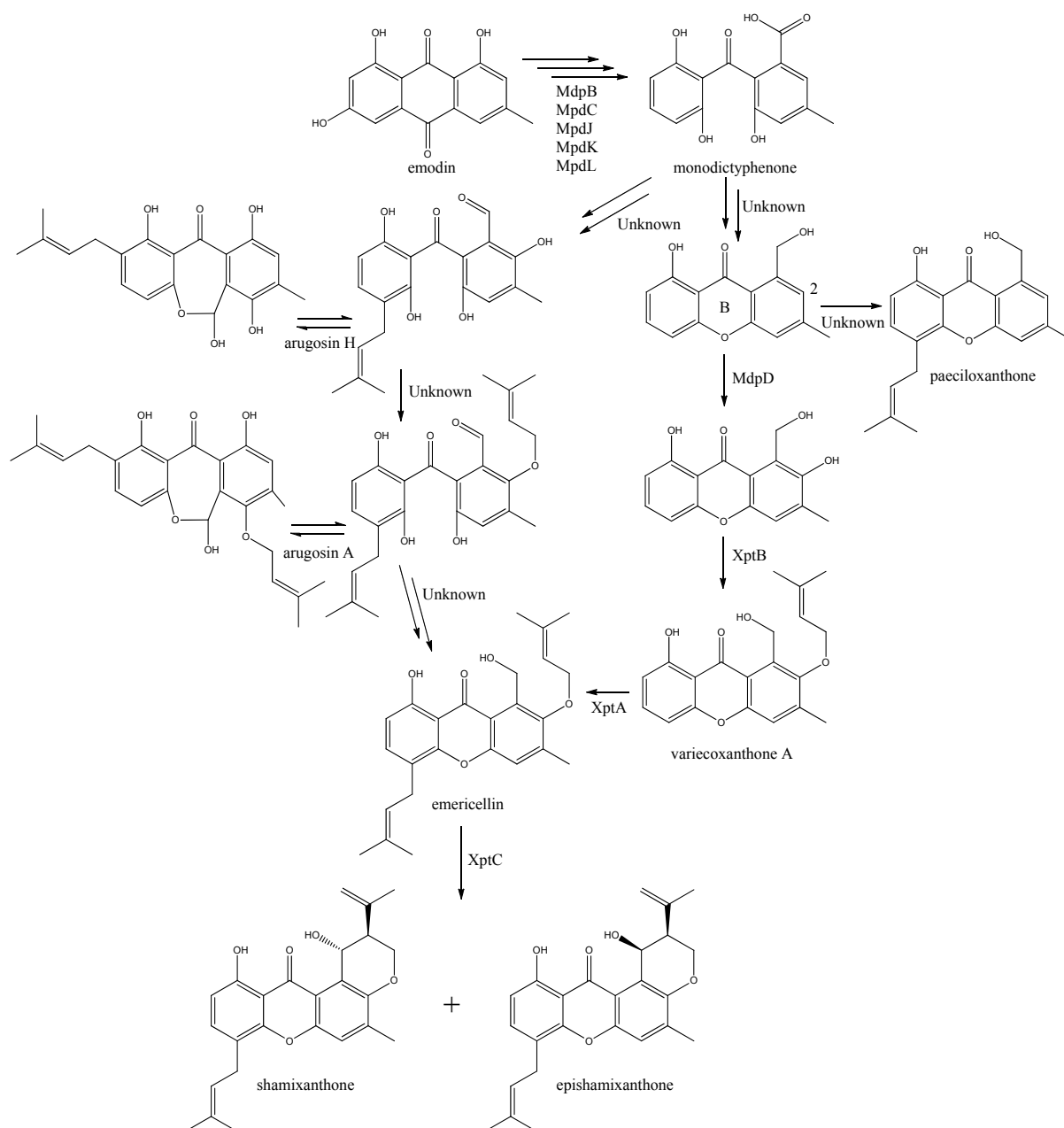


The first stable intermediate following emodin towards the prenyl-xanthenes is monodictyphenone [97,98], and gene-deletion studies points to at least the five following enzymes are involved; a dehydratase (MdpB), a ketoreductase (MdpC), a glutathione S transferase (MdpJ), an oxidoreductase (MdpK), and a Baeyer-Villiger oxidase (MdpL) [97]. The mechanism has been proposed to be analogous to the conversion of versicolorin A to demethylsterigmatocystin which is known to proceed through oxidation-reduction-oxidation catalyzed by a cytochrome P450 monooxygenase (VerA) and a ketoreductase (StcU) [45,65,110]. However, none of the above mentioned Mdp enzymes appear to be homologous to VerA, and the role of the individual enzymes has not been investigated further [97].

The biosynthesis of the six monodictyphenone derived metabolites is based on hydroxylation (MdpD), C-prenylation (XptA), O-prenylation (XptB), and carboxylic acid reduction (unidentified enzyme) [11,98]. Central in the pathway is the hydroxylation of C2 in monodictyphenone accompanied by reduction of the carboxylic acid. The carboxylic acid is suggested by Sanchez and co-workers [98] to be reduced to a hydroxy group, the B-ring is closed by dehydration and the

intermediate is O-prenylated at C2 to yield variecoxanthone A, which in turn is C-prenylated to emericillin (Figure 10). The final known step in prenyl-xanthone biosynthesis gives rise to the stereoisomers shamixanthone and epishamixanthone and is catalyzed by XptC [98]. Alternatively, Nielsen and co-workers include synthesis of arugosins by partially reducing the carboxylic acid to an aldehyde, followed by C-prenylation, yielding arugosin H and O-prenylation to give arugosin A. Subsequent reduction of the aldehyde to a hydroxyl group, and ring closure by dehydration then gives emericillin and shamixanthones [11].

Figure 10. Suggested biosynthesis of the shamixanthones from emodin. Multiple arrows indicate that the number of enzymatic steps are unknown.



9. Orsellinic Acid

In addition to the *mdp* cluster, the loss of CclA also led to the discovery of another gene cluster driven by an NR-PKS [96]. Two PK products, the cathepcin K inhibitors F-9775A and F-9775B, first isolated from *Paecilomyces carneus* [111], were detected and mapped to AN7909. Following this discovery, Schroeckh and co-workers [112] found the primary metabolite from AN7909 (*orsA*) to be orsellinic acid, an archetypal metabolite [113]. Moreover, the metabolite lecanoric acid typically found in lichens and produced by mycobiots such as *Umbilicaria antarctica* [114] was linked to OrsA [112]. Following the initial observations, the number of detected metabolites from the *orsA* gene cluster has expanded to gerfelin, a C10-deoxy-gerfelin, diorcinol, orcinol, cordyol C, and violaceol I and II [115,116]. The biosynthetic activities of the *orsA* cluster are as yet not elucidated, and this illustrates the need for applying different eloquent strategies to trigger production of these metabolites.

The deficiency in methylation of H3K4 in the *cclAΔ* strain resulted in activation of both *mdp* and *ors* gene clusters. Expression analysis revealed that the annotated ORFs from AN7909-AN7915 were possible cluster members, hereby indicating candidates for a gene cluster [96]. The *ors* gene cluster, *orsB-orsE*, was identified by Schroeckh and co-workers [112] as four additional ORFs spanning AN7911-AN7914, which was confirmed by Sanchez *et al.* [115], who deleted all genes from AN7901 to AN7915. Interestingly, the neighbor PKS to the *ors* locus, AN7903, was deleted by Nielsen *et al.* [11] and the resulting strain failed to produce F-9775A and B like AN7909Δ under the conditions tested. Schroeckh and co-workers [112] defined *orsA-E* using gene-expression analysis through both an *Aspergillus* secondary metabolism array (ASMA) and relative expression analysis in quantitative reverse-transcriptase PCR (qRT-PCR). The induction of *orsA* was achieved by co-cultivating with a soil bacterium, *Streptomyces rapamycinicus* (initially named *S. hygroscopicus*) and extracting mRNA from the fungus [112]. This response on SM level was further investigated by Nützmann and co-workers [117]. Since the loss of H3K4 methylation induced gene expression in *ors* locus [96], the rationale was that the transcriptional activation of silent secondary-metabolism genes by acetylation of lysines on histone tails, especially H3K9, is equally important and the search for histone acetyl transferases (HATs) in the genome sequence was commenced [117]. Forty HATs were found and deleted, and only four proved to be essential. Of the 36 deletions of nonessential HATs in *A. nidulans*, the deletions of *genE* and *adaA*, both essential core parts of the multi-subunit Saga/Ada complex, an important complex for HAT activity in *A. nidulans*, significantly lowered the *ors* transcripts investigated [117]. Thus, Saga/Ada plays a role in the response to *S. rapamycinicus* and loss of this complex downregulated orsellinic acid metabolites, as well as sterigmatocystin, terrequinone, and penicillin [117].

Four additional orsellinic acid derived compounds were found in a defect COP9 signalosome (CSN) mutant strain [116]. The multiunit CSN complex is found in higher eukaryotes, albeit with different functional roles depending on the tissues. In *A. nidulans* the CSN is required for fruiting body formation and is not essential for asexual growth. By deleting *csnE*, orcinol, cordyol C, and violaceol I+II were produced, and the genes *orsA-orsE* were shown to be differentially expressed [116]. The link of the violaceol metabolites to *ors* was confirmed by Nielsen and co-workers [11] who applied an OSMAC strategy on their reference strain and compared this to their deletion library.

Very little is known about the biosynthesis of the metabolites of the *ors* locus. One acetyl-CoA and three malonyl-CoA units can yield a C8 aldol intermediate, and as proposed by Nielsen *et al.* [11], this can lead to the tetraketide orsellinic acid through loss of water and enolization and to the C7 compound orcinol by decarboxylation and enolization. Oxidation of orcinol in the para position then leads to 5-methyl-benzene-1,2,3-triol which is believed to either dimerize with the loss of water to give violaceol I and II, Figure 11, or to give F-9775A+B, Figure 12, in an unknown series of synthesis steps. Another outcome is the formation of lecanoric acid by dimerization of orsellinic acid. Though the steps in the pathway have been hypothesized, most steps are not accounted for. It has been reported that OrsA having the domains SAT-KS-AT-PT-ACP-TE is responsible solely to form orsellinic acid. OrsA, OrsB, and OrsC seem to be sole responsible for F-9775A+B formation. Moreover, it has been shown that gerfelin and a C10-deoxy derivative of gerfelin accumulate in *orsBΔ*, whereas diorcinol was found in high amounts in the *orsCΔ* strain [115]. Gerfelin, C10-deoxy gerfelin and cordyol C are all dimers built up of two of the three suggested monomer units, orsellinic acid, orcinol and 5-methylbenzene-1,2,3-triol.

Recently Scherlach and co-workers [118] continuously cultivated *A. nidulans* under nitrogen-limitation and carbon-limitation. At nitrogen limiting conditions in continuous cultivations two novel products, denoted as spiroanthrones, were found. They could not be detected at batch cultivation. The metabolites were based on anthraquinone and orsellinic acid derived phenols. The induced expression of both *mdpG* and *orsA* confirmed increased activity under the N-limiting continuous cultivation conditions.

10. Austinol and Dehydroaustinol

The meroterpenoids austinol and the related compound dehydroaustin were first isolated from *A. ustus* by Simpson and co-workers in 1982 [119], where the structure of austinol was elucidated by ¹H and ¹³C NMR. Austinol and dehydroaustinol are just two examples out of many meroterpenoids that are derived from 3,5-dimethyl orsellinic acid as presented in the excellent review by Geris and Simpson [120].

The two austinols were detected for the first time in *A. nidulans* four years ago [121], where it was further substantiated that austinol was indeed of partly PK origin. Deletion of the phosphopantetheinyl transferase (PPTase) *cfwA/npgA* in *A. nidulans* resulted in a strain that among many other compounds did not produce austinol and dehydroaustinol [121]. The PPTase is responsible for attaching the phosphopantetheine moiety to the acyl carrier domain of the PKSs and NRPSs, thus it is an activator of the enzyme complexes. Hence, the abolition of austinol and dehydroaustinol production in the PPTase deficient strain strongly suggests a PK origin of these compounds.

In 2011, Nielsen and co-workers [11] discovered the PKS responsible for synthesis of the PK part of austinol and dehydroaustinol in *A. nidulans*. A deletion library of all 32 putative PKS genes in *A. nidulans* was created and screened using an OSMAC approach [122] to enable activation of different clusters on different media. One single strain deleted in AN8383 (*ausA*) failed to produce austinol and dehydroaustinol [11]. This discovery was supported by the introduction of a point mutation at the phosphopantetheine attachment site, thus abolishing activation of the enzyme by the PPTase to ensure the loss of austinols was not an indirect effect e.g., at chromatin level. The *ausAΔ*

strain was complemented by re-introducing the *ausA* gene into the deletion strain. Introducing the gene under control of the inducible *alcA* promoter revealed that 3,5-dimethyl orsellinic acid (3,5-MOA) was indeed the precursor for austinol and dehydroaustinol in *A. nidulans*, as shown experimentally via labeling studies by Simpson and co-workers 20 years earlier [11,123,124].

Figure 11. Proposed biosynthesis of orsellinic acid and its derivatives of orsellinic acid. The enzymes that catalyze the individual reactions in the biosynthesis of the metabolites are so far unknown and biosynthesis is proposed based on the observed metabolites.

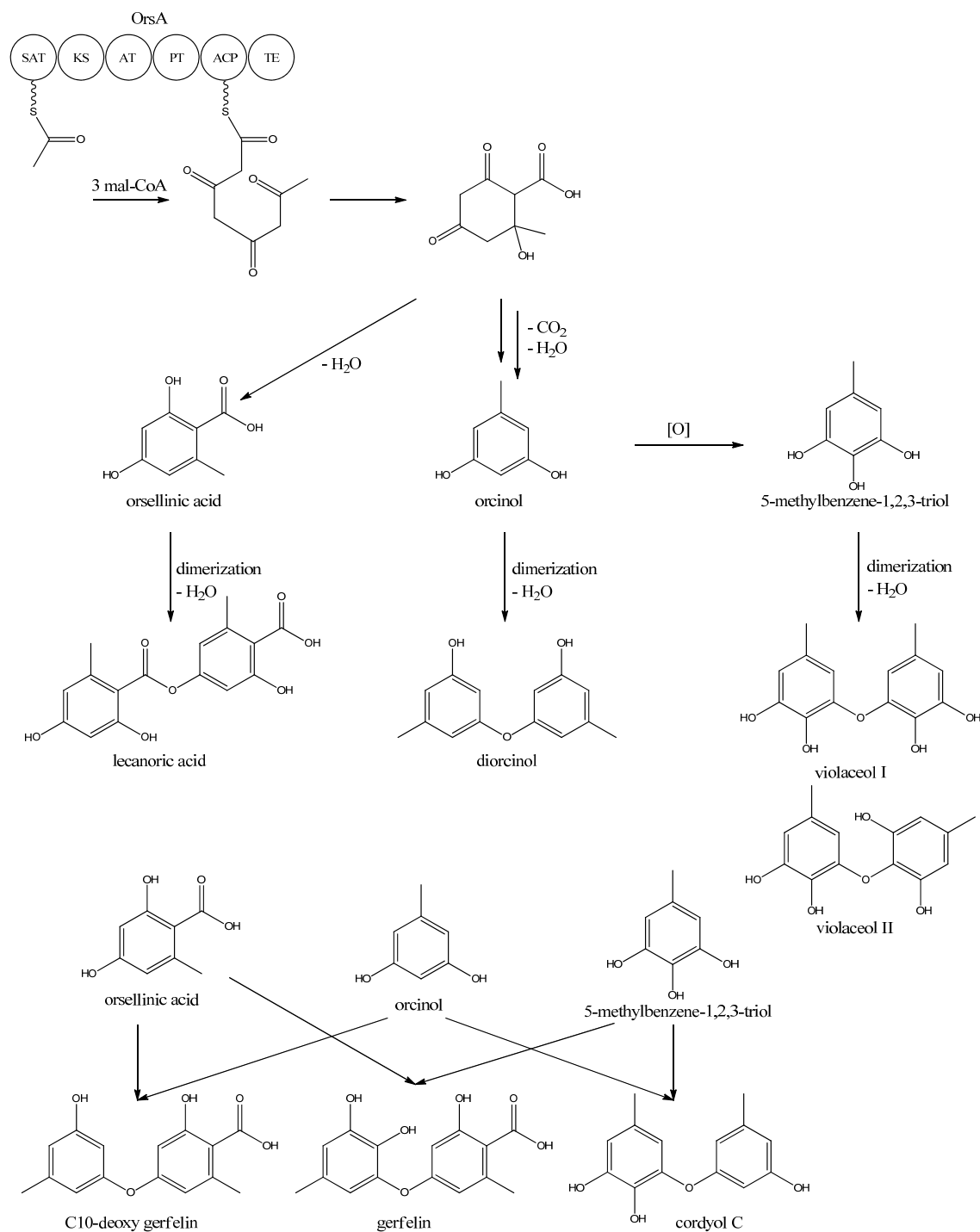
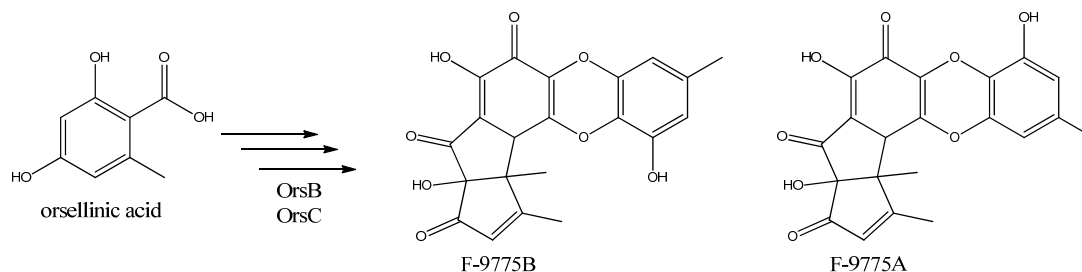


Figure 12. The proposed enzymes involved in the biosynthesis of F-9775A and F-9775B of the figure here.



A more detailed understanding of the biosynthesis of the austinols has not been established yet. However, it is well established that 3,5-dimethyl orsellinic acid is synthesized from condensation of an acetyl-CoA unit with 3 malonyl-CoA units to form the PK backbone, which is methylated twice, catalyzed by AusA. The PK part is then alkylated with farnesyl pyrophosphate to form a transient intermediate (Figure 13), which can then act as a precursor for several similar meroterpenoids, such as andibenins, austin, berkeleyones and andrastins [120]. Recently we tentatively identified neo-austin and austinolide in *A. nidulans* extracts by LC-MS analysis (unpublished data), which makes us hypothesize that the biosynthesis towards the austinols involves (i) oxidation and acyl shift in the D ring (ii) lactonization from the substituent groups of the D ring, a Baeyer-Villiger type oxidation and 1,2 alkyl shift in the A ring to give neo-austin. Neo-austin is subsequently oxidized in the D ring by another Baeyer-Villiger type oxidation to give austinolide that upon further oxidation and ring condensation leads to austinol and dehydroaustinol, Figure 13.

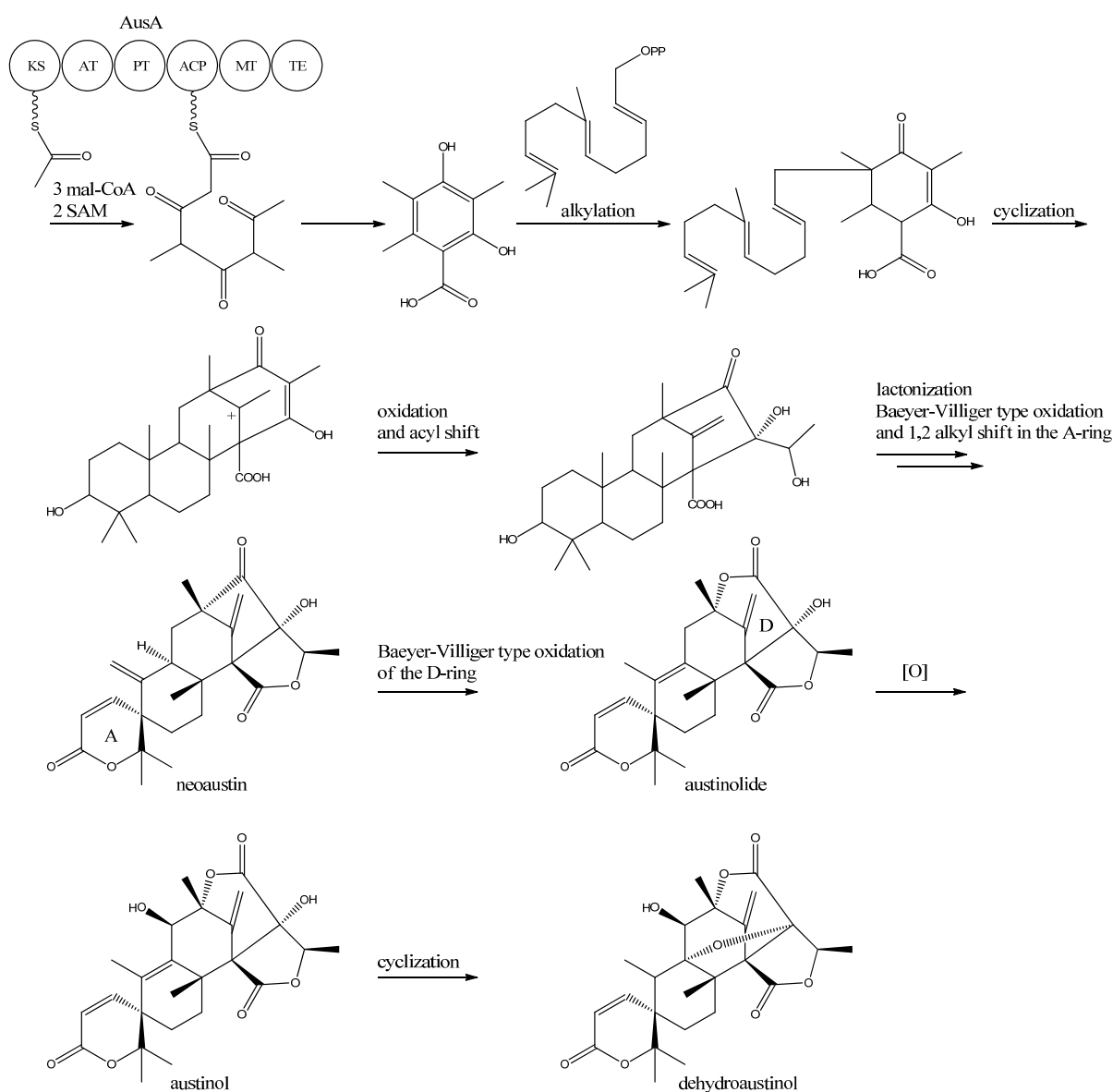
11. Concluding Remarks

Secondary metabolism represents chemical diversity and span in biological functionality to the extreme. As shown above, individual compound classes can even form hybrid molecules to other compound classes. There is a high commercial interest in discovery and utilization of SMs in general as drugs or additives, or to avoid mycotoxins in food and feeds. Mapping PK biosynthesis to genes in *A. nidulans* as presented in this review involves great complexity. One challenge is to find and activate the genes required to produce the compounds. As shown, it takes in-depth understanding of fungal biology, nutrient sensing, chromatin remodeling, as well as analysis on all levels from DNA to metabolites to unveil cryptic gene clusters and their products. Moreover, the majority of the pathways described in this review have been elucidated in a relatively short time span. This has been possible due to bioinformatics. The availability of the genome sequence, as well as resources and tools in, e.g., *Aspergillus* Comparative Database (ACD), *Aspergillus* Genome Database (AspGD), and the Central *Aspergillus* REsource (CADRE) have been key aids to perform the extensive genome mining.

The ability to predict enzymatic function based on gene sequences has proven fruitful in characterization of secondary metabolism, since this has revealed the location of e.g., PKS genes. Additionally, the presence of gene cluster specific TFs was utilized in activation of silent SM clusters both in the case of aspyridones and asperfuranone. The asperfuranone biosynthesis is moreover an example of a previously unknown compound, with a potential to be a novel drug, has been found in a

well-known filamentous fungus. Another case of a metabolite with attractive properties is emodin, which has been known for more than a century, but just recently had the biosynthetic machinery uncovered in *A. nidulans*. Both asperthecin and the emericellamides were firstly discovered and isolated from less well described aspergilli, however, after the compounds were observed in *A. nidulans*, the candidates for responsible genes in biosynthesis were found in a few months. The study of conidial pigment biosynthesis in *A. nidulans* has contributed to our basic understanding of fungal development and PKS organization, and provided researchers with an easy assayable marker system for genetic studies. Interestingly, the structure of the final pigment(s) still remains unknown after more than 70 years of research, underlining the difficulties in elucidating structures of highly polymerized PKS.

Figure 13. Proposed biosynthesis of austinol and dehydroaustinol. All genes in the biosynthesis of austinol and dehydroaustinol, except AusA, are unknown.



Epigenetic regulation through chromatin remodeling has shown to be involved in activation of several SM gene clusters. Conserved chromatin remodeling factors have influenced both local activation of some SM clusters and more global responses within the genome. Gene clusters producing sterigmatocystin, orsellinic acid, emodins, and austinols have shown to respond to specific factors. The presence of SM producing genes outside gene clusters, e.g., in prenyl-xanthone production, is probably more common than observed so far. Moreover, cross-talk between pathways is frequently observed, as more pathways become known. This can open a discussion whether common pools of intermediates or enzymes can exist. In addition, the compartmentalization of SM production is an area to be explored. Furthermore, controlling compartmentalization of production as well as secretion will influence yields and downstream applications which are important factors for exploiting SM production commercially. Existing compounds can be modified genetically to add/remove chemical groups on existing drugs, mix moieties from different SM classes or species by for example domain swapping, and to considerably increase/abolish a specific production.

Altogether the recent uncovering of secondary metabolism in *A. nidulans* is an illustrative example of strong interdisciplinary efforts requiring strong expertise within chemistry, biology, microbiology, molecular genetics, protein chemistry, computer science, and engineering. Ultimately, the efforts described in this review can form the basis for uncovering of the specific biological roles of the chemical arsenal in the fungus.

Acknowledgments

This work was supported by the Danish Research Agency for Technology and Production grant # 09-064967, FOBI, DTU, and Novozymes. Additionally, we thank Maria Månsson for critical proof reading of the manuscript and for suggestions.

Conflict of Interest

The authors declare no conflict of interest.

References and Notes

1. Pontecorvo, G.; Roper, J.A.; Hemmons, L.M.; MacDonald, K.D.; Bufton, A.W.J. The genetics of *Aspergillus nidulans*. *Adv. Genet. Incomp. Mol. Genet. Med.* **1953**, *5*, 141–238.
2. Newman, D.J.; Cragg, G.M. Natural products as sources of new drugs over the last 25 years. *J. Nat. Prod.* **2007**, *70*, 461–477.
3. Crawford, J.M.; Townsend, C.A. New insights into the formation of fungal aromatic polyketides. *Nat. Rev. Microbiol.* **2010**, *8*, 879–889.
4. Hertweck, C. The biosynthetic logic of polyketide diversity. *Angew. Chem. Int. Ed.* **2009**, *48*, 4688–4716.
5. Cox, R.J. Polyketides, proteins and genes in fungi: Programmed nano-machines begin to reveal their secrets. *Org. Biomol. Chem.* **2007**, *5*, 2010–2026.
6. Bingle, L.E.H.; Simpson, T.J.; Lazarus, C.M. Ketosynthase domain probes identify two subclasses of fungal polyketide synthase genes. *Fun. Gen. Biol.* **1999**, *26*, 209–223.

7. Crawford, J.M.; Vagstad, A.L.; Ehrlich, K.C.; Townsend, C.A. Starter unit specificity directs genome mining of polyketide synthase pathways in fungi. *Bioorg. Chem.* **2008**, *36*, 16–22.
8. Crawford, J.M.; Thomas, P.M.; Scheerer, J.R.; Vagstad, A.L.; Kelleher, N.L.; Townsend, C.A. Deconstruction of iterative multidomain polyketide synthase function. *Science* **2008**, *320*, 243–246.
9. Yadav, G.; Gokhale, R.S.; Mohanty, D. Towards prediction of metabolic products of polyketide synthases: An *in silico* analysis. *PLoS Comp. Biol.* **2009**, *5*, e1000351:1–e1000351:14.
10. Du, L.; Lou, L. PKS and NRPS release mechanisms. *Nat. Prod. Rep.* **2010**, *27*, 255–278.
11. Nielsen, M.L.; Nielsen, J.B.; Rank, C.; Klejnstrup, M.L.; Holm, D.K.; Brogaard, K.H.; Hansen, B.G.; Frisvad, J.C.; Larsen, T.O.; Mortensen, U.H. A genome-wide polyketide synthase deletion library uncovers novel genetic links to polyketides and meroterpenoids in *Aspergillus nidulans*. *FEMS Microbiol. Lett.* **2011**, *321*, 157–166.
12. Chung, Y.S.; Kim, J.M.; Han, D.M.; Chae, K.S.; Jahng, K.Y. Ultrastructure of the cell wall of a null pigmentation mutant, *npgA1*, in *Aspergillus nidulans*. *J. Microbiol.* **2003**, *41*, 224–231.
13. Jahn, B.; Boukhallouk, F.; Lotz, J.; Langfelder, K.; Wanner, G.; Brakhage, A.A. Interaction of human phagocytes with pigmentless *Aspergillus* conidia. *Infect. Immun.* **2000**, *68*, 3736–3739.
14. Wright, P.J.; Pateman, J.A. Ultraviolet-light sensitive mutants of *Aspergillus nidulans*. *Mutation Res.* **1970**, *9*, 579–587.
15. Yuill, E. Two new *Aspergillus* mutants. *J. Bot.* **1939**, *77*, 174–175.
16. Clutterbuck, A.J. Absence of laccase from yellow-spored mutants of *Aspergillus nidulans*. *J. Gen. Microbiol.* **1972**, *70*, 423–435.
17. Agnihotri, V.P. Role of trace elements in the growth and morphology of five ascosporic *Aspergillus* species. *Can. J. Bot.* **1967**, *45*, 73–79.
18. O'Hara, E.B.; Timberlake, W.E. Molecular characterization of the *Aspergillus nidulans* *yA* locus. *Genetics* **1989**, *121*, 249–254.
19. Mayorga, M.E.; Timberlake, W.E. Isolation and molecular characterization of the *Aspergillus nidulans* *wA* gene. *Genetics* **1990**, *126*, 73–79.
20. Marshall, M.A.; Timberlake, W.E. *Aspergillus nidulans* *wetA* activates spore-specific gene expression. *Mol. Cell. Biol.* **1991**, *11*, 55–62.
21. Adams, T.H.; Wieser, J.K.; Yu, J.H. Asexual sporulation in *Aspergillus nidulans*. *Microbiol. Mol. Biol. Rev.* **1998**, *62*, 35–54.
22. Mayorga, M.E.; Timberlake, W.E. The developmentally regulated *Aspergillus nidulans* *wA* gene encodes a polypeptide homologous to polyketide and fatty acid synthases. *Mol. Gen. Genet.* **1992**, *235*, 205–212.
23. Watanabe, A.; Ono, Y.; Fujii, I.; Sankawa, U.; Mayorga, M.E.; Timberlake, W.E.; Ebizuka, Y. Product identification of polyketide synthase coded by *Aspergillus nidulans* *wA* gene. *Tetrahedron Lett.* **1998**, *39*, 7733–7736.
24. Fujii, I.; Watanabe, A.; Sankawa, U.; Ebizuka, Y. Identification of Claisen cyclase domain in fungal polyketide synthase *WA*, a naphthopyrone synthase of *Aspergillus nidulans*. *Chem. Biol.* **2001**, *8*, 189–197.
25. Watanabe, A.; Fujii, I.; Sankawa, U.; Mayorga, M.E.; Timberlake, W.E.; Ebizuka, Y. Re-identification of *Aspergillus nidulans* *wA* gene to code for a polyketide synthase of naphthopyrone. *Tetrahedron Lett.* **1999**, *40*, 91–94.

26. Watanabe, A.; Ebizuka, Y. A novel hexaketide naphthalene synthesized by a chimeric polyketide synthase composed of fungal pentaketide and heptaketide synthase. *Tetrahedron Lett.* **2002**, *43*, 843–846.
27. Nekam, L.; Polgar, P. The inhibitory effect of a mold upon *Staphylococcus*. *Urol. Cutan. Rev.* **1948**, *52*, 372–373.
28. Hatsuda, Y.; Kuyama, S. Studies on the metabolic products of *Aspergillus versicolor*. Part 1. Cultivation of *Aspergillus versicolor*, isolation and purification of metabolic products. *J. Agric. Chem. Soc. Jpn.* **1954**, *28*, 989–991.
29. Hatsuda, Y.; Kuyama, S.; Terashima, N. Studies on the metabolic products of *Aspergillus versicolor*. Part 2. Physical and chemical properties and the chemical structure of sterigmatocystin. *J. Agric. Chem. Soc. Jpn.* **1954**, *28*, 992–998.
30. Bullock, E.; Underwood, J.G.V.; Roberts, J.C. Studies in mycological chemistry. 11. Structure of isosterigmatocystin and an amended structure for sterigmatocystin. *J. Chem. Soc.* **1962**, *10*, 4179–4183.
31. Holker, J.S.E.; Mulheirn, L.J. Biosynthesis of sterigmatocystin. *Chem. Comm.* **1968**, *24*, 1574–1579.
32. Brechbühler, S.; Buchi, G.; Milne, G. Absolute configuration of aflatoxins. *J. Org. Chem.* **1967**, 2641–2642.
33. Fukuyama, K.; Hamada, K.; Tsukihara, T.; Katsube, Y.; Hamasaki, T.; Hatsuda, Y. Crystal structures of sterigmatocystin and *O*-methylsterigmatocystin, metabolites of genus *Aspergillus*. *Bull. Chem. Soc. Jpn.* **1976**, *49I*, 1153–1154.
34. Fukuyama, K.; Tsukihara, T.; Katsube, Y.; Tanaka, N.; Hamasaki, T.; Hatsuda, Y. Crystal and molecular structure of *p*-bromobenzoate of sterigmatocystin. *Bull. Chem. Soc. Jpn.* **1975**, *48*, 1980–1983.
35. Blount, W.P. Turkey “X” disease. *Turkey* **1961**, *61*, 55–58.
36. Terao, K. Sterigmatocystin—A masked potent carcinogenic mycotoxin. *J. Toxicol. Toxin Rev.* **1983**, *2*, 77–110.
37. Bennett, J.W.; Klinch, M. Mycotoxins. *Clin. Microbiol. Rev.* **2003**, *16*, 497–516.
38. Bhatnager, D.; Yu, J.; Ehrlich, K.C. Toxins of filamentous fungi. *Chem. Immunol.* **2002**, *81*, 167–206.
39. Yu, J.; Chang, P.K.; Ehrlich, K.C.; Cary, J.W.; Bhatnagar, D.; Cleveland, T.E.; Payne, G.A.; Linz, J.E.; Woloshuk, C.P.; Bennett, J.W. Clustered pathway genes in aflatoxin biosynthesis. *Appl. Environ. Microbiol.* **2004**, *70*, 1253–1262.
40. Yabe, K.; Nakajima, H. Enzyme reactions and genes in aflatoxin biosynthesis. *Appl. Microbiol. Biotechnol.* **2004**, *64*, 745–755.
41. Brown, D.W.; Yu, J.H.; Kelkar, H.S.; Fernandes, M.; Nesbitt, T.C.; Keller, N.P.; Adams, T.H.; Leonard, T.J. Twenty-five coregulated transcripts define a sterigmatocystin gene cluster in *Aspergillus nidulans*. *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 1418–1422.
42. Yu, J.H.; Leonard, T.J. Sterigmatocystin biosynthesis in *Aspergillus nidulans* requires a novel type I polyketide synthase. *J. Bacteriol.* **1995**, *177*, 4792–4800.
43. McDonald, T.; Hammond, T.; Noordermeer, D.; Zhang, Y.Q.; Keller, N. The sterigmatocystin cluster revisited: lessons from a genetic model. In *Food Science and Technology/Aflatoxin and Food Safety*; Abbas, H.K., Ed.; CRC Press: Boca Raton, FL, USA, 2005; pp. 117–136.

44. Ehrlich, K.C.; Li, P.; Scharfenstein, L.; Chang, P.K. HypC, the anthrone oxidase involved in aflatoxin biosynthesis. *Appl. Environ. Microbiol.* **2010**, *76*, 3374–3377.
45. Henry, K.M.; Townsend, C.A. Ordering the reductive and cytochrome P450 oxidative steps in demethylsterigmatocystin formation yields general insights into the biosynthesis of aflatoxin and related fungal metabolites. *J. Am. Chem. Soc.* **2005**, *127*, 3724–3733.
46. Ehrlich, K.C.; Montalbano, B.; Boué, S.M.; Bhatnagar, D. An aflatoxin biosynthesis cluster gene encodes a novel oxidase required for conversion of versicolorin A to sterigmatocystin. *Appl. Environ. Microbiol.* **2005**, *71*, 8963–8965.
47. Cary, J.W.; Ehrlich, K.C.; Bland, J.M.; Montalbano, B.G. The aflatoxin biosynthesis cluster gene, *aflX*, encodes an oxidoreductase involved in conversion of versicolorin A to demethylsterigmatocystin. *Appl. Environ. Microbiol.* **2006**, *72*, 1096–1101.
48. Yu, J.H.; Butchko, R.A.E.; Fernandes, M.; Keller, N.P.; Leonard, T.J.; Adams, T.H. Conservation of structure and function of the aflatoxin regulatory gene *aflR* from *Aspergillus nidulans* and *A. flavus*. *Curr. Genet.* **1996**, *29*, 549–555.
49. Meyers, D.M.; Obrian, G.; Du, W.L.; Bhatnagar, D.; Payne, G.A. Characterization of *aflJ*, a gene required for conversion of pathway intermediates to aflatoxin. *Appl. Environ. Microbiol.* **1998**, *64*, 3713–3717.
50. Bok, J.W.; Keller, N.P. LeaA, a regulator of secondary metabolism in *Aspergillus* spp. *Eukaryot. Cell* **2004**, *3*, 527–535.
51. Bayram, Ö.; Krappmann, S.; Ni, M.; Bok, J.W.; Helmstaedt, K.; Valerius, O.; Braus-Stromeier, S.; Kwon, N.J.; Keller, N.P. VelB/VeA/LaeA complex coordinates light signal with fungal development and secondary metabolism. *Science* **2008**, *320*, 1504–1506.
52. Shwab, E.K.; Bok, J.W.; Tribus, M.; Galehr, J.; Graessle, S.; Keller, N.P. Histone deacetylase activity regulates chemical diversity in *Aspergillus*. *Eukaryot. Cell* **2007**, *6*, 1656–1664.
53. Fisch, K.M.; Gillaspay, A.F.; Gipson, M.; Henrikson, J.C.; Hoover, A.R.; Jackson, L.; Najar, F.Z.; Wägele, H.; Cichewicz, R.H. Chemical induction of silent biosynthetic pathway transcription in *Aspergillus niger*. *J. Ind. Microbiol. Biotechnol.* **2009**, *36*, 1199–1213.
54. Watanabe, C.M.H.; Townsend, C.A. Initial characterization of a type I fatty acid synthase and polyketide synthase multienzyme complex NorS in the biosynthesis of aflatoxin B₁. *Chem. Biol.* **2002**, *9*, 981–988.
55. Yabe, K.; Nakamura, Y.; Nakajima, H.; Ando, Y.; Hamasaki, T. Enzymatic conversion of norsolorinic acid to averufin in aflatoxin biosynthesis. *Appl. Environ. Microbiol.* **1991**, *57*, 1340–1345.
56. Trail, F.; Chang, P.K.; Cary, J.; Linz, J.E. Structural and functional analysis of the *nor-1* gene involved in the biosynthesis of aflatoxins by *Aspergillus parasiticus*. *Appl. Environ. Microbiol.* **1994**, *60*, 4078–4085.
57. Keller, N.P.; Watanabe, C.M.H.; Kelkar, H.S.; Adams, T.H.; Townsend, C.A. Requirement of monooxygenase-mediated steps for sterigmatocystin biosynthesis by *Aspergillus nidulans*. *Appl. Environ. Microbiol.* **2000**, *66*, 359–362.
58. Yabe, L.; Matsuyama, Y.; Ando, Y.; Nakajima, H.; Hamasaki, T. Stereochemistry during aflatoxin biosynthesis: Conversion of norsolorinic acid to averufin. *Appl. Environ. Microbiol.* **1993**, *59*, 2486–2492.

59. Chang, P.K.; Yu, J.; Ehrlich, K.C.; Boue, S.M.; Montalbano, B.G.; Bhatnagar, D.; Cleveland, T.E. *adhA* in *Aspergillus parasiticus* is involved in conversion of 5'-hydroxyaverantin to averufin. *Appl. Environ. Microbiol.* **2000**, *66*, 4715–4719.
60. Sakuno, E.; Yabe, K.; Nakajima, H. Involvement of two cytosolic enzymes and a novel intermediate 5'-oxoaverantin in the pathway from 5'-hydroxyaverantin to averufin in aflatoxin biosynthesis. *Appl. Environ. Microbiol.* **2003**, *69*, 6418–6426.
61. Chang, P.K.; Yabe, K.; Yu, J. The *Aspergillus parasiticus estA*-encoded esterase converts versiconal hemiacetal acetate to versiconal and versiconol acetate to versiconol in aflatoxin biosynthesis. *Appl. Environ. Microbiol.* **2004**, *70*, 3593–3599.
62. Silva, J.C.; Minto, R.E.; Barry, C.E., III; Holland, K.A.; Townsend, C.A. Isolation and characterization of the versicolorin B synthase gene from *Aspergillus parasiticus*. *J. Biol. Chem.* **1996**, *271*, 13600–13608.
63. Kelkar, H.S.; Skloss, T.W.; Haw, J.F.; Keller, N.P.; Adams, T.H. *Aspergillus nidulans stcL* encodes a putative cytochrome P-450 monooxygenase required for bisfuran desaturation during aflatoxin/sterigmatocystin biosynthesis. *J. Biol. Chem.* **1997**, *272*, 1589–1594.
64. Keller, N.P.; Kantz, N.J.; Adams, T.H. *Aspergillus nidulans verA* is required for production of the mycotoxin sterigmatocystin. *Appl. Environ. Microbiol.* **1994**, *60*, 1444–1450.
65. Keller, N.P.; Segner, S.; Bhatnagar, D.; Adams, T.H. *stcS*, a putative P-450 monooxygenase, is required for the conversion of versicolorin A to sterigmatocystin in *Aspergillus nidulans*. *Appl. Environ. Microbiol.* **1995**, *61*, 3628–3632.
66. Kelkar, H.S.; Keller, N.P.; Adams, T.H. *Aspergillus nidulans stcP* encodes an *O*-methyltransferase that is required for sterigmatocystin biosynthesis. *Appl. Environ. Microbiol.* **1996**, *62*, 4296–4298.
67. Slot, J.C.; Rokas, A. Horizontal transfer of a large and highly toxic secondary metabolic gene cluster between fungi. *Curr. Biol.* **2011**, *21*, 134–139.
68. Bergmann, S.; Schumann, J.; Scherlach, K.; Lange, C.; Brakhage, A.A.; Hertweck, C. Genomics-driven discovery of PKS-NRPS hybrid metabolites from *Aspergillus nidulans*. *Nat. Chem. Biol.* **2007**, *3*, 213–217.
69. Chiang, Y.M.; Szewczyk, E.; Nayak, T.; Davidson, A.D.; Sanchez, J.F.; Lo, H.C.; Ho, W.Y.; Simityan, H.; Kuo, E.; Praseuth, A.; *et al.* Molecular genetic mining of the *Aspergillus* secondary metabolome: Discovery of the emericellamide biosynthetic pathway. *Chem. Biol.* **2008**, *15*, 527–532.
70. Sims, J.W.; Fillmore, J.P.; Warner, D.D.; Schmidt, E.W. Equisetin biosynthesis in *Fusarium heterosporum*. *Chem. Commun.* **2005**, 186–188.
71. Song, Z.; Cox, R.J.; Lazarus, C.M.; Simpson, T.J. Fusarin C biosynthesis in *Fusarium moniliforme* and *Fusarium venenatum*. *ChemBioChem* **2004**, *5*, 1196–1203.
72. Liu, X.; Walsh, C.T. Cyclopiazonic acid biosynthesis in *Aspergillus* sp.: Characterization of a reductase-like R* domain in cyclopiazonate synthetase that forms and releases *cyclo*-acetoacetyl-L-tryptophan. *Biochemistry* **2009**, *48*, 8746–8757.
73. Xu, W.; Cai, X.; Jung, M.E.; Tang, Y. Analysis of intact and dissected fungal polyketide synthase-nonribosomal peptide synthetase *in vitro* and in *Saccharomyces cerevisiae*. *J. Am. Chem. Soc.* **2010**, *132*, 13604–13607.

74. Halo, L.M.; Marshall, J.W.; Yakasai, A.A.; Song, Z.; Butts, C.P.; Crump, M.P.; Heneghan, M.; Bailey, A.M.; Simpson, T.J.; Lazarus, C.M.; Cox, R.J. Authentic heterologous expression of the tenellin iterative polyketide synthase nonribosomal peptide synthetase requires coexpression with an enoyl reductase. *ChemBioChem* **2008**, *9*, 585–594.
75. Eley, K.L.; Halo, L.M.; Song, Z.; Powles, H.; Cox, R.J.; Bailey, A.M.; Lazarus, C.M.; Simpson, T.J. Biosynthesis of the 2-pyridone tenellin in the insect pathogenic fungus *Beauveria bassiana*. *ChemBioChem* **2007**, *8*, 289–297.
76. Fujita, Y.; Oguri, H.; Oikawa, H. Biosynthetic studies on the antibiotics PF1140: a novel pathway for a 2-pyridone framework. *Tetrahedron Lett.* **2005**, *46*, 5885–5888.
77. McInnes, A.G.; Smith, D.G.; Wat, C.K. Tenellin and Bassianin, Metabolites of *Beauveria* Species. Structure elucidation with ¹⁵N- and doubly ¹³C-enriched compounds using ¹³C nuclear magnetic resonance spectroscopy. *J. Chem. Soc. Chem. Commun.* **1974**, *8*, 281–282.
78. Halo, L.M.; Heneghan, M.N.; Yakasai, A.A.; Song, Z.; Williams, K.; Bailey, A.M.; Cox, R.J.; Lazarus, C.M.; Simpson, T.J. Late stage oxidations during the biosynthesis of the 2-pyridone tenellin in the entomopathogenic fungus *Beauveria bassiana*. *J. Am. Chem. Soc.* **2008**, *130*, 17988–17996.
79. Oh, D.C.; Kauffman, C.A.; Jensen, P.R.; Fenical, W. Induced production of emericellamides A and B from the marine-derived fungus *Emericella* sp. in competing co-culture. *J. Nat. Prod.* **2007**, *70*, 515–520.
80. Howard, B.H.; Raistrick, H. Studies in the biochemistry of micro-organism. 94. The colouring matters of species in the *Aspergillus nidulans* group. Part I. Asperthecin, a crystalline colouring matter of *Aspergillus quadrilineatus* Thom & Raper. *Biochem. J.* **1955**, *59*, 475–484.
81. Birkinshaw, J.H.; Gourlay, R. Studies in the biochemistry of micro-organisms. 109. The structure of asperthecin. *Biochem. J.* **1961**, *81*, 618–622.
82. Neelakantan, S.; Pocker, A.; Raistrick, H. Studies in the biochemistry of micro-organisms. 101. The colouring matters of species in the *Aspergillus nidulans* group. Part 2. Further observations on the structure of asperthecin. *Biochem. J.* **1957**, *66*, 234–237.
83. Szewczyk, E.; Chiang, Y.M.; Oakley, C.E.; Davidson, A.D.; Wang, C.C.C.; Oakley, B.R. Identification and characterization of the asperthecin gene cluster of *Aspergillus nidulans*. *Appl. Environ. Microbiol.* **2008**, *74*, 7607–7612.
84. Wong, K.H.; Todd, R.B.; Oakley, B.R.; Oakley, C.E.; Hynes, M.J.; Davis, M.A. Sumoylation in *Aspergillus nidulans*: *sumO* inactivation, overexpression and live-cell imaging. *Fungal Genet. Biol.* **2008**, *45*, 728–737.
85. Awakawa, T.; Yokota, K.; Funa, N.; Doi, F.; Mori, N.; Watanabe, H.; Horinouchi, S. Physically discrete β -lactamase-type thioesterase catalyzes product release in atrochrysone synthesis by iterative type I polyketide synthase. *Chem. Biol.* **2009**, *16*, 613–623.
86. Chiang, Y.M.; Oakley, B.R.; Keller, N.P.; Wang, C.C.C. Unraveling polyketide synthesis in members of the genus *Aspergillus*. *Appl. Microbiol. Biotechnol.* **2010**, *86*, 1719–1736.
87. Li, Y.; Chooi, Y.H.; Sheng, Y.; Valentine, J.S.; Tang, Y. Comparative characterization of fungal anthracenone and naphthacenedione biosynthetic pathways reveals an α -hydroxylation-dependent claisen-like cyclization catalyzed by a dimanganese thioesterase. *J. Am. Chem. Soc.* **2011**, *133*, 15773–15785.

88. Li, Y.; Xu, W.; Tang, Y. Classification, prediction, and verification of the regioselectivity of fungal polyketide synthase product template domains. *J. Biol. Chem.* **2010**, *285*, 22764–22773.
89. Chiang, Y.M.; Szewczyk, E.; Davidson, A.D.; Keller, N.; Oakley, B.R.; Wang, C.C.C. A gene cluster containing two fungal polyketide synthases encodes the biosynthetic pathway for a polyketide, asperfuranone, in *Aspergillus nidulans*. *J. Am. Chem. Soc.* **2009**, *131*, 2965–2970.
90. Wang, C.C.C.; Chiang, Y.M.; Praseuth, M.B.; Kuo, P.L.; Liang, H.L.; Hsu, Y.L. Asperfuranone from *Aspergillus nidulans* inhibits proliferation of human non-small cell lung cancer A549 cells via blocking cell cycle progression and inducing apoptosis. *Basic Clin. Pharmacol. Toxicol.* **2010**, *107*, 583–589.
91. Matsuzaki, K.; Tahara, H.; Inokoshi, J.; Tanaka, H.; Masuma, R.; Omura, S. New brominated and halogen-less derivatives and structure-activity relationship of azaphilones inhibiting gp120-CD4 binding. *J. Antibiot.* **1998**, *51*, 1004–1011.
92. Bailey, A.M.; Cox, R.J.; Harley, K.; Lazarus, C.M.; Simpson, T.J.; Skellam, E. Characterisation of 3-methylorcinolaldehyde synthase (MOS) in *Acremonium strictum*: First observation of a reductive release mechanism during polyketide biosynthesis. *Chem. Commun.* **2007**, *39*, 4053–4055.
93. Bergmann, S.; Funk, A.N.; Scherlach, K.; Schroeckh, V.; Shelest, E.; Horn, U.; Hertweck, C.; Brakhage, A.A. Activation of a silent fungal polyketide biosynthesis pathway through regulatory cross talk with a cryptic nonribosomal peptide synthetase gene cluster. *Appl. Environ. Microbiol.* **2010**, *76*, 8143–8149.
94. Liu, T.; Chiang, Y.M.; Somoza, A.D.; Oakley, B.R.; Wang, C.C.C. Engineering of an “unnatural” natural product by swapping polyketide synthase domains in *Aspergillus nidulans*. *J. Am. Chem. Soc.* **2011**, *133*, 13314–13316.
95. Lu, P.; Zhang, A.; Dennis, L.M.; Dhal-Roshak, A.M.; Xia, Y.Q.; Arison, B.; An, Z.; Tkacz, J.S. A gene (*pks2*) encoding a putative 6-methylsalicylic acid synthase from *Glarea lozoyensis*. *Mol. Gen. Genomics* **2005**, *273*, 207–216.
96. Bok, J.W.; Chiang, Y.M.; Szewczyk, E.; Reyes-Dominguez, Y.; Davidson, A.D.; Sanchez, J.F.; Lo, H.C.; Watanabe, K.; Strauss, J.; Oakley, B.R.; *et al.* Chromatin-level regulation of biosynthetic gene clusters. *Nat. Chem. Biol.* **2009**, *5*, 462–464.
97. Chiang, Y.M.; Szewczyk, E.; Davidson, A.D.; Entwistle, R.; Keller, N.P.; Wang, C.C.C.; Oakley, B.R. Characterization of the *Aspergillus nidulans* monodictyphenone gene cluster. *Appl. Environ. Microbiol.* **2010**, *76*, 2067–2074.
98. Sanchez, J.F.; Entwistle, R.; Hung, J.H.; Yaegashi, J.; Jain, S.; Chiang, Y.M.; Wang, C.C.C.; Oakley, B.R. Genome-based deletion analysis reveals the prenyl xanthone biosynthesis pathway in *Aspergillus nidulans*. *J. Am. Chem. Soc.* **2011**, *133*, 4010–4017.
99. Jowett, H.A.D.; Potter, C.E. CXXVII—The constitution of chrysophante acid and of emodin. *J. Chem. Soc. Trans.* **1903**, *83*, 1327–1334.
100. Beal, G.D.; Okey, R. The qualitative identification of the drugs containing emodin. *J. Am. Chem. Soc.* **1917**, *39*, 716–725.
101. Gatenbeck, S. Incorporation of labelled acetate in emodin in *Penicillium islandicum*. *Acta Chem. Scand.* **1958**, *12*, 1211–1214.

102. Basu, S.; Ghosh, A.; Hazra, B. Evaluation of the antibacterial activity of *Ventilago madraspatana* Gaertn., *Rubia cordifolia* Linn. and *Lantana camara* Linn.: Isolation of emodin and physcion as active antibacterial agents. *Phytother. Res.* **2005**, *19*, 888–894.
103. Chen, J.; Zhang, L.; Zhang, Y.; Zhang, H.; Du, J.; Ding, J.; Guo, Y.; Jiang, H.; Shen, X. Emodin targets the β -hydroxyacyl carrier protein dehydratase from *Helicobacter pylori*: Enzymatic inhibition assay with crystal structural and thermodynamic characterization. *BMC Microbiol.* **2009**, *9*, 91:1–91:12.
104. Huang, Q.; Lu, G.; Shen, H.M.; Chung, M.C.M.; Ong, C.N. Anti-cancer properties of anthraquinones from rhubarb. *Med. Res. Rev.* **2007**, *27*, 609–630.
105. Srinivas, G.; Babykutty, S.; Sathiadevan, P.P.; Srinivas, P. Molecular mechanism of emodin action: Transition from laxative ingredient to an antitumor agent. *Med. Res. Rev.* **2007**, *27*, 591–608.
106. Zargar, B.A.; Masoodi, M.H.; Ahmed, B.; Ganie, S.A. Phytoconstituents and therapeutic uses of *Rheum emodi* wall. Ex Meissn. *Food Chem.* **2011**, *128*, 585–589.
107. Chen, Z.G.; Fujii, I.; Ebizuka, Y.; Sankawa, U. Purification and characterization of emodinanthrone oxygenase from *Aspergillus terreus*. *Phytochemistry* **1995**, *38*, 299–305.
108. Huang, K.X.; Fujii, I.; Ebizuka, Y.; Gomi, K.; Sankawa, U. Molecular cloning and heterologous expression of the gene encoding dihydroemodin oxidase, a multicopper blue enzyme from *Aspergillus terreus*. *J. Biol. Chem.* **1995**, *270*, 21495–21502.
109. Couch, R.D.; Gaucher, G.M. Rational elimination of *Aspergillus terreus* sulochrin production. *J. Biotech.* **2004**, *108*, 171–178.
110. Henry, K.M.; Townsend, C.A. Synthesis and fate of *o*-carboxybenzophenones in the biosynthesis of aflatoxin. *J. Am. Chem. Soc.* **2005**, *127*, 3300–3309.
111. Sato, S.; Morishita, T.; Hosoya, T.; Ishikawa, Y. Novel pentacyclic compounds, F-9775A and F-9775B, their manufacture with *Paecilomyces carneus*, and their use for treatment of osteoporosis. Japan Pat. JP11001480A 19990106, 1999.
112. Schroeckh, V.; Scherlach, K.; Nützmann, H.W.; Shelest, E.; Schmidt-Heck, W.; Schuemann, J.; Martin, K.; Hertweck, C.; Brakhage, A.A. Intimate bacterial-fungal interaction triggers biosynthesis of archetypal polyketides in *Aspergillus nidulans*. *Proc. Natl. Acad. Soc. USA* **2009**, *106*, 14558–14563.
113. Gaucher, G.M.; Shepherd, M.G. Isolation of orsellinic acid synthase. *Biochem. Biophys. Res. Commun.* **1968**, *32*, 664–671.
114. Leo, H.; Yamamoto, Y.; Kim, J.A.; Jung, J.S.; Koh, Y.J.; Hur, J.S. Lecanoric acid, a secondary lichen substance with antioxidant properties from *Umbilicaria antarctica* in maritime Antarctica (King George Island). *Polar Biol.* **2009**, *32*, 1033–1040.
115. Sanchez, J.F.; Chiang, Y.M.; Szewczyk, E.; Davidson, A.D.; Ahuja, M.; Oakley, C.E.; Bok, J.W.; Keller, N.P.; Oakley, B.R.; Wang, C.C.C. Molecular genetic analysis of the orsellinic acid/F9775 gene cluster of *Aspergillus nidulans*. *Mol. Biosyst.* **2010**, *6*, 587–593.
116. Nahlik, K.; Dumkow, M.; Bayram, Ö.; Helmstaedt, K.; Busch, S.; Valerius, O.; Gerke, J.; Hoppert, M.; Schwier, E.; Opitz, L.; et al. The COP9 signalosome mediates transcriptional and metabolic response to hormones, oxidative stress production and cell wall rearrangement during fungal development. *Mol. Microbiol.* **2010**, *78*, 964–979.

117. Nützmann, H.W.; Reyes-Doninguez, Y.; Scherlach, K.; Schroeckh, V.; Horn, F.; Gacek, A.; Schümann, J.; Hertweck, C.; Strauss, J.; Brakhage, A.A. Bacteria-induced natural product formation in the fungus *Aspergillus nidulans* requires Saga/Ada-mediated histone acetylation. *Proc. Natl. Acad. Soc. USA* **2011**, *108*, 14282–14287.
118. Scherlach, K.; Sarkar, A.; Schroeckh, V.; Dahse, H.M.; Roth, M.; Brakhage, A.A.; Horn, W.; Hertweck, C. Two induced fungal polyketide pathways converge into antoproliferative spiroanthrones. *ChemBioChem* **2011**, *12*, 1836–1839.
119. Simpson, T.J.; Stenzel, D.J.; Bartlett, A.J.; O'Brien, E.; Holker, J.S.E. Studies on fungal metabolites. Part 3. ^{13}C -NMR spectral and structural studies on Austin and new related meroterpenoids from *Aspergillus ustus*, *Aspergillus varicolor*, *Penicillium diversum*. *J. Chem. Soc. Perkin Trans. I* **1982**, *11*, 2687–2692.
120. Geris, R.; Simpson, T.J. Meroterpenoids produced by fungi. *Nat. Prod. Rep.* **2009**, *26*, 1063–1094.
121. Márquez-Fernández, O.; Trigos, Á.; Ramos-Balderas, J.L.; Viniegra-González, G.; Deising, H.B.; Aguirre, J. Phosphopantetheinyl transferase CfwA/NpgA is required for *Aspergillus nidulans* secondary metabolism and asexual development. *Eukaryot. Cell* **2007**, *6*, 710–720.
122. Bode, H.B.; Bethe, B.; Höfs, R.; Zeeck, A. Big effects from small changes: possible ways to explore nature's chemical diversity. *ChemBioChem* **2002**, *3*, 619–627.
123. Simpson, T.J.; Stenzel, D.J. Biosynthesis of austin, a polyketide-terpenoid metabolite of *Aspergillus ustus*. *JCS Chem. Commun.* **1981**, *20*, 1042–1043.
124. Scott, F.E.; Simpson, T.J.; Trimble, L.A.; Vederas, J.C. Biosynthesis of the meroterpenoid austin by *Aspergillus ustus*: Synthesis and incorporation of ^{13}C , ^{18}O -labelled ethyl 3,5-dimethylorsellinate. *J.C.S. Chem. Comm.* **1986**, *3*, 214–215.

© 2012 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/3.0/>).

Chapter 6

Identification of novel secondary metabolites by over-expression of phosphatases and transcription factors from *A. nidulans*

6.1 Introduction

A. nidulans is a well-characterized filamentous fungus with respect to genetics and secondary metabolism. The genome sequencing of *A. nidulans* [Galagan et al., 2005] revealed that its genome contained at least 32 polyketide synthases (PKS), which are large multidomain enzymes that are responsible for synthesis of polyketides (PKs). However, this number greatly exceeds the number of known PKs from *A. nidulans* and although quite a few *A. nidulans* polyketides are actually characterized they are rarely produced under the same circumstances, further complicating the investigation of the polyketome and appertaining PKSs. For instance, the huge impact on the polyketide production by the composition of growth medium is shown in Figure 6.1 [Klejnstrup, 2012]. Additionally, pH, water activity, temperature, atmosphere, light etc. also influence the regulation of secondary metabolism [Samson et al., 2010].

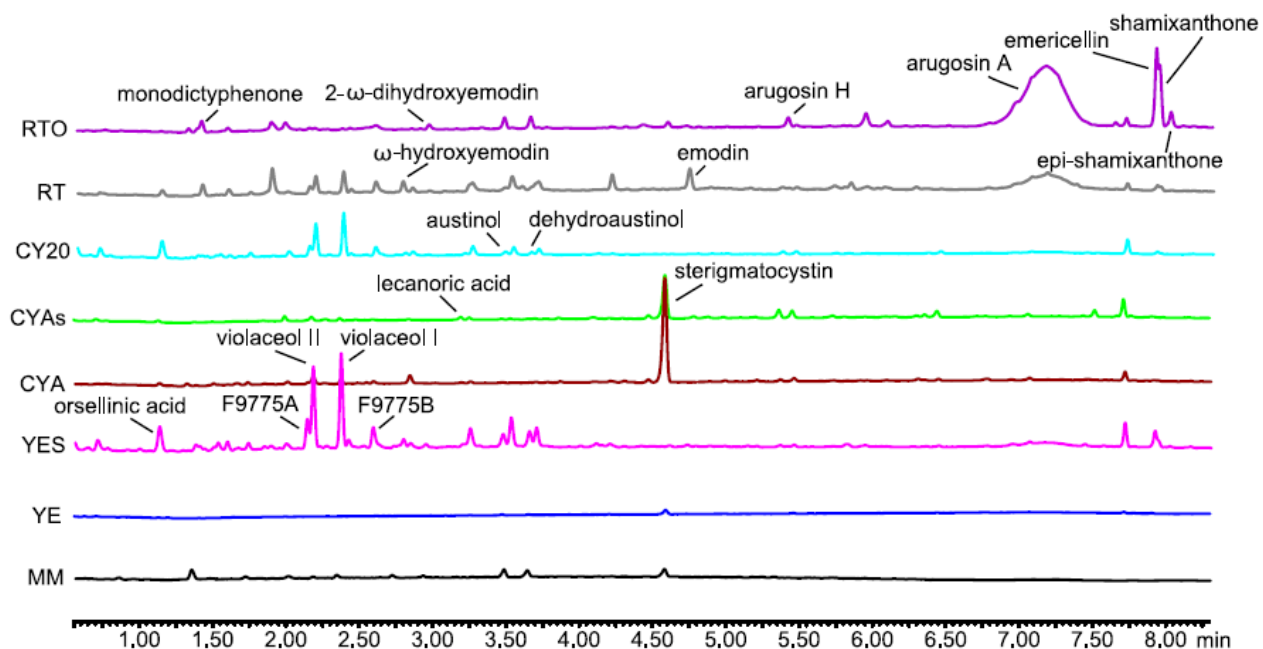


Figure 6.1: Impact of growth medium composition on the *A. nidulans* polyketome. Figure from [Klejnstrup, 2012].

Many different efforts have been done in order to illuminate the polyketome of *A. nidulans* and link genes to their corresponding PK, including heterologous expression of PKSs [Watanabe et al., 1998, 1999, Clutterbuck, 1972], over-expression of local transcription factors (TFs) [Bergmann et al., 2007, Chiang et al., 2009], gene deletion or inactivation [Chiang et al., 2008, Nielsen et al., 2011, Yaegashi et al., 2013, Yu and Leonard, 1995, Brown et al., 1996], global chromatin regulation [Bok et al., 2009], perturbation of sumoylation [Szewczyk et al., 2008], and co-cultivation with other organisms [Schroeckh et al., 2009]. In total, 10 PKSs have been linked to nine PKs in *A. nidulans*, suggesting that there may still be more than twenty PK awaiting discovery.

As many creative and original approaches have already been applied for triggering production of PKs, the solution for activating the remaining silent clusters may not be obvious or trivial. Having developed a system for high-throughput gene expression in *A. nidulans*, I assessed the possibility to systematically over-express *A. nidulans* genes and examine the metabolite profiles of the resulting strains. As *A. nidulans* contains 9,541 putative protein encoding genes [Galagan et al., 2005], over-expression of each individual gene would entail a tremendous workload. Sopko and co-workers have previously performed systematic over-expression of all genes in *S. cerevisiae*, and reported that the cells were more sensitive to perturbations of pathways that are defined by proteins involved in transport, cell-cycle regulation, transcriptional activation, and signaling e.g. kinases and phosphatases [Sopko et al., 2006].

In a very recent study Yaegashi and co-workers have shown that deletion of a kinase indeed activated a silent cluster in *A. nidulans* [Yaegashi et al., 2013, De Souza et al., 2013], and results from previous

studies on *A. nidulans* secondary metabolism have shown that over-expression of TFs also could lead to activation of silent gene clusters. *A. nidulans* contains 28 putative phosphatases, which are the protein kinase counterparts that catalyze protein dephosphorylation [Son and Osmani, 2009], and 372 putative TFs [Galagan et al., 2005]. As a proof-of-concept study I decided to over-express all 28 phosphatases and all 52 annotated putative TFs on chromosome I, comprising 80 genes in total. If successful, novel PKs could easily be linked to their corresponding PKSs in combination with the genome-wide PKS deletion library constructed in [Nielsen et al., 2011] (Chapter 4).

6.2 Results and discussion

This study was used as a test case for high-throughput fungal strain construction, and comprised very little manual work. PCR primers were designed using the PHUSER program (Chapter 7), PCR reactions were set up using a Hamilton STAR liquid handler, PCR fragment lengths were determined via capillary electrophoresis, assessed via a novel computer program, VirtualPCR (Ali Altintas, not published), and inserted directly into the cloning vector pDH124 without purification (Figure 6.2). Likewise, plasmids were linearized by digestion and transformed directly into *A. nidulans* without purification. The cloning vector contained the O7 promoter (TetON/OFF), which allows tunable expression by different concentrations of doxycycline (DOX). The *A. nidulans* recipient strain allowed for background-free insertion of expression constructs into IS1, enabling extremely fast verification of correct transformants.

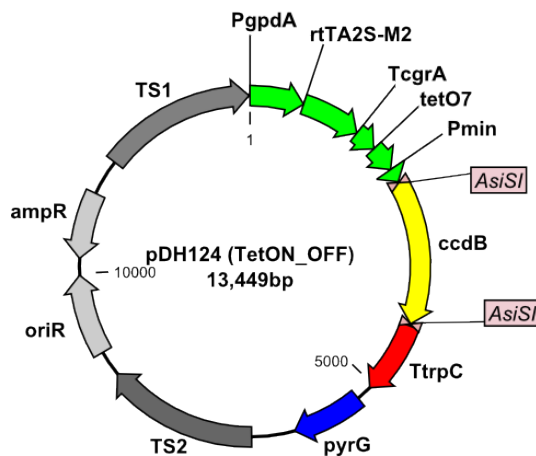


Figure 6.2: Background-free cloning vector pDH124 employing TetON/OFF.

6.2.1 Over-expression of phosphatases affects growth rate and morphology

The phosphatase expression strains were constructed as part of my external research stay at Ohio State University, and for some reason I had problems with preparation of protoplasts, why only half of the constructs provided transformants. However, 13 correct phosphatase expression strains were constructed using the

background-free integration system. As I had no prior experience with the use of the TetON/OFF system, I used four different concentrations of DOX for assessment of the morphologies. The morphologies of the strains grown on MM+DOX are shown in Figure 6.3, large pictures are found in Appendix B.

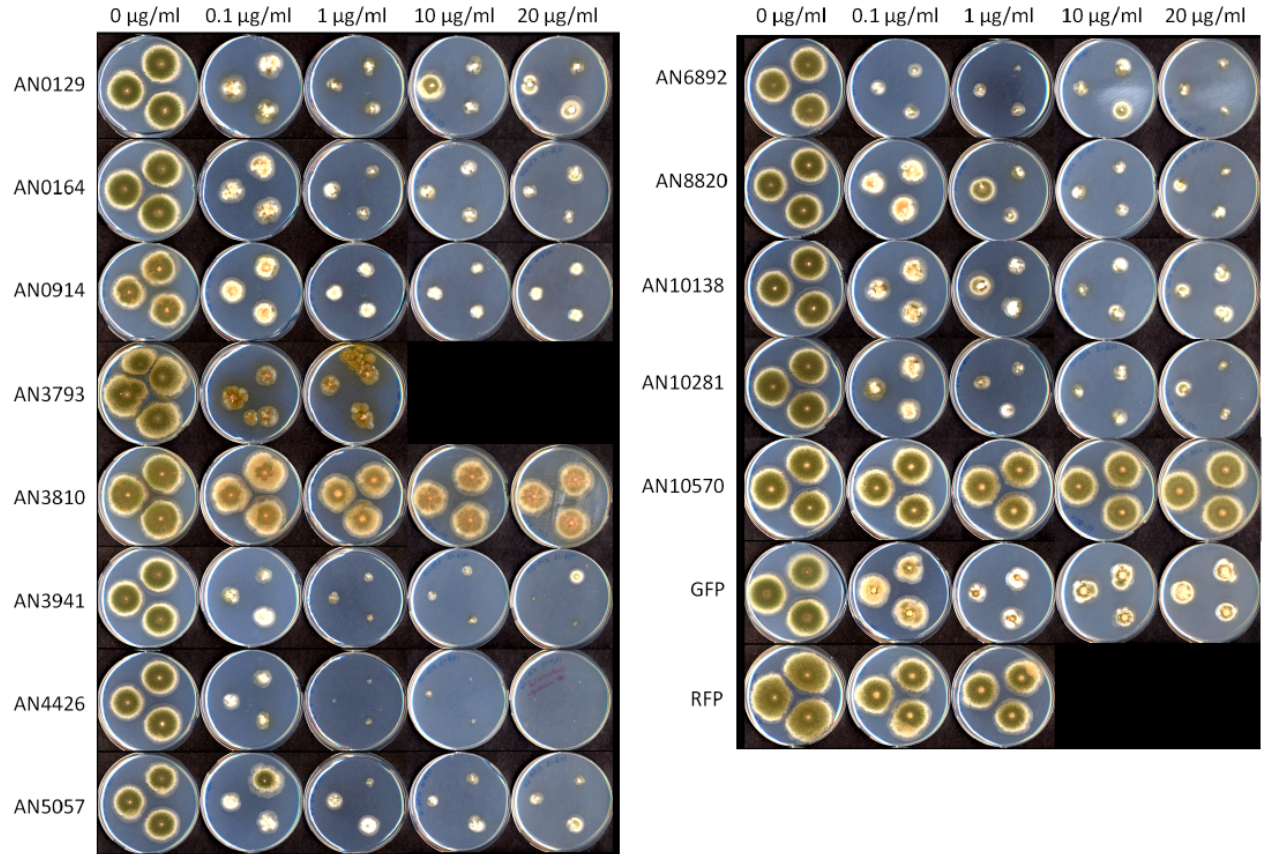


Figure 6.3: Macromorphologies of phosphatase over-expression strains cultivated at four different concentrations of DOX. GFP = reference with *gfp* inserted, RFP = reference with *rfp* inserted.

Two strains seem to be only little affected by over-expression of the phosphatases: AN3810 and AN10570. In contrast, the remaining 13 strains, including the GFP- and RFP-expressing references, had markedly reduced growth. The colony diameters are listed in Table 6.2.1, and plotted in Figure 6.4. Intriguingly, the strains seem to form wild-type appearing wedges from the slow-growing colonies, which also results in large error-bars in the plot for some of the strains. The wedges may be formed as a result of a spontaneous loop-out event in an effort to get rid of the clearly toxic gene. The construct does not contain an apparent direct repeat, so it would be interesting to sequence the strains and examine what has happened.

Table 6.1: Average colony diameters (mm) of phosphatase over-expression strains.

Strain	0 $\mu\text{g/ml}$ DOX	0.1 $\mu\text{g/ml}$ DOX	1.0 $\mu\text{g/ml}$ DOX	10 $\mu\text{g/ml}$ DOX	20 $\mu\text{g/ml}$ DOX
RFP	15.4 \pm 0.4	13.0 \pm 0.9	11.7 \pm 0.6	N/A	N/A
GFP	12.8 \pm 0.0	9.8 \pm 0.9	6.7 \pm 0.6	6.9 \pm 0.4	6.9 \pm 0.4
AN0129	13.0 \pm 0.4	8.1 \pm 0.7	4.3 \pm 0.4	6.1 \pm 3.3	5.7 \pm 1.9
AN0164 (<i>ppg1A</i>)	13.0 \pm 0.4	8.0 \pm 0.7	4.4 \pm 1.1	4.8 \pm 0.4	4.3 \pm 1.3
AN0914	11.3 \pm 0.4	7.6 \pm 0.4	5.2 \pm 0.4	4.3 \pm 0.4	4.4 \pm 0.0
AN3793 (<i>ppzA</i>)	13.7 \pm 0.9	6.5 \pm 0.7	6.1 \pm 0.0	N/A	N/A
AN3810	13.5 \pm 0.4	12.8 \pm 0.6	11.5 \pm 0.4	11.7 \pm 0.0	11.1 \pm 0.0
AN3941 (<i>nimT</i>)	12.0 \pm 0.4	5.7 \pm 0.9	2.8 \pm 1.1	2.6 \pm 0.4	2.2 \pm 2.2
AN4426	11.7 \pm 0.6	5.2 \pm 0.4	1.3 \pm 0.4	1.1 \pm 0.6	0.6 \pm 0.0
AN5057 (<i>cdcA</i>)	11.1 \pm 0.6	7.8 \pm 3.3	4.8 \pm 2.6	4.1 \pm 0.9	4.3 \pm 1.3
AN6892 (<i>ptcA</i>)	11.7 \pm 0.0	4.6 \pm 0.4	3.7 \pm 1.5	5.7 \pm 0.9	3.1 \pm 0.9
AN8820 (<i>cnaA</i>)	11.5 \pm 0.4	8.5 \pm 0.4	5.9 \pm 2.4	4.1 \pm 0.4	3.9 \pm 1.1
AN10138	12.4 \pm 0.4	8.1 \pm 0.4	6.7 \pm 2.2	4.4 \pm 0.6	5.2 \pm 0.4
AN10281	13.0 \pm 0.4	7.6 \pm 0.4	4.4 \pm 0.6	4.3 \pm 0.4	3.7 \pm 1.9
AN10570 (<i>ltpA</i>)	13.3 \pm 0.6	12.4 \pm 0.4	12.6 \pm 0.4	12.4 \pm 0.4	13.1 \pm 0.4

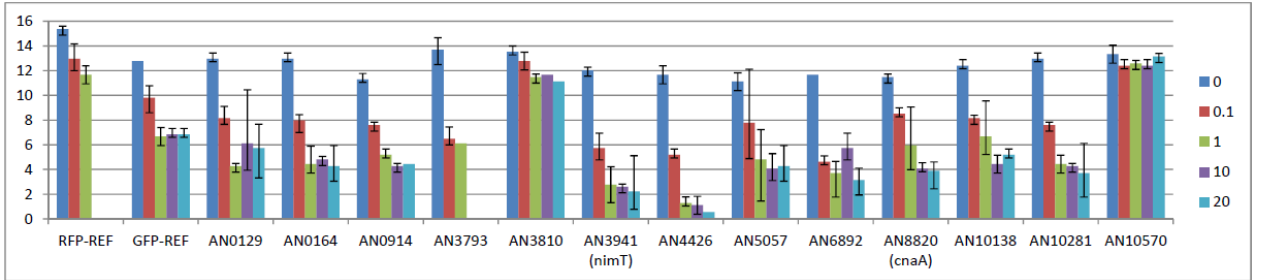


Figure 6.4: Histogram of colony diameters of phosphatase over-expression strains. Series values represent DOX concentration ($\mu\text{g/ml}$).

6.2.2 Over-expression of TFs results in several new phenotypes

In addition to the phosphatase strains, 39 correct TF expression strains were constructed. As the morphologies of the phosphatase strains did not change much between 1 $\mu\text{g/ml}$ and 20 $\mu\text{g/ml}$ DOX, the TF expression strains were cultivated on MM+0.1 $\mu\text{g/ml}$ DOX and MM+1 $\mu\text{g/ml}$ DOX for assessing the effect of gene over-expression on morphology and growth rate. An overview of the strains are shown in Figure 6.5, and

large pictures are found in Appendix B.

Surprisingly, the TF over-expression strains did not seem severely impaired, not did they display the same tendency to form wedges as was observed for the phosphatase over-expression strains. In contrast, the TF strains looked mostly healthy, although somewhat growth impaired. Diameters of the colonies are listed in Table 6.2.2. Additionally, the TF strains displayed very diverse morphologies including changes in conidiation levels and production of pigments. An interesting observation is that the AN10763 strain, in contrast to the other strains, is greener and more healthy-looking on 1.0 $\mu\text{g}/\text{ml}$ DOX than on 0.1 $\mu\text{g}/\text{ml}$ DOX, which may indicate a higher degree of conidiation on 1.0 $\mu\text{g}/\text{ml}$ DOX. A puzzling observation for many of the strains is that the colonies differ a lot in size i.e. the growth rates seem to vary for individuals of the same strain.

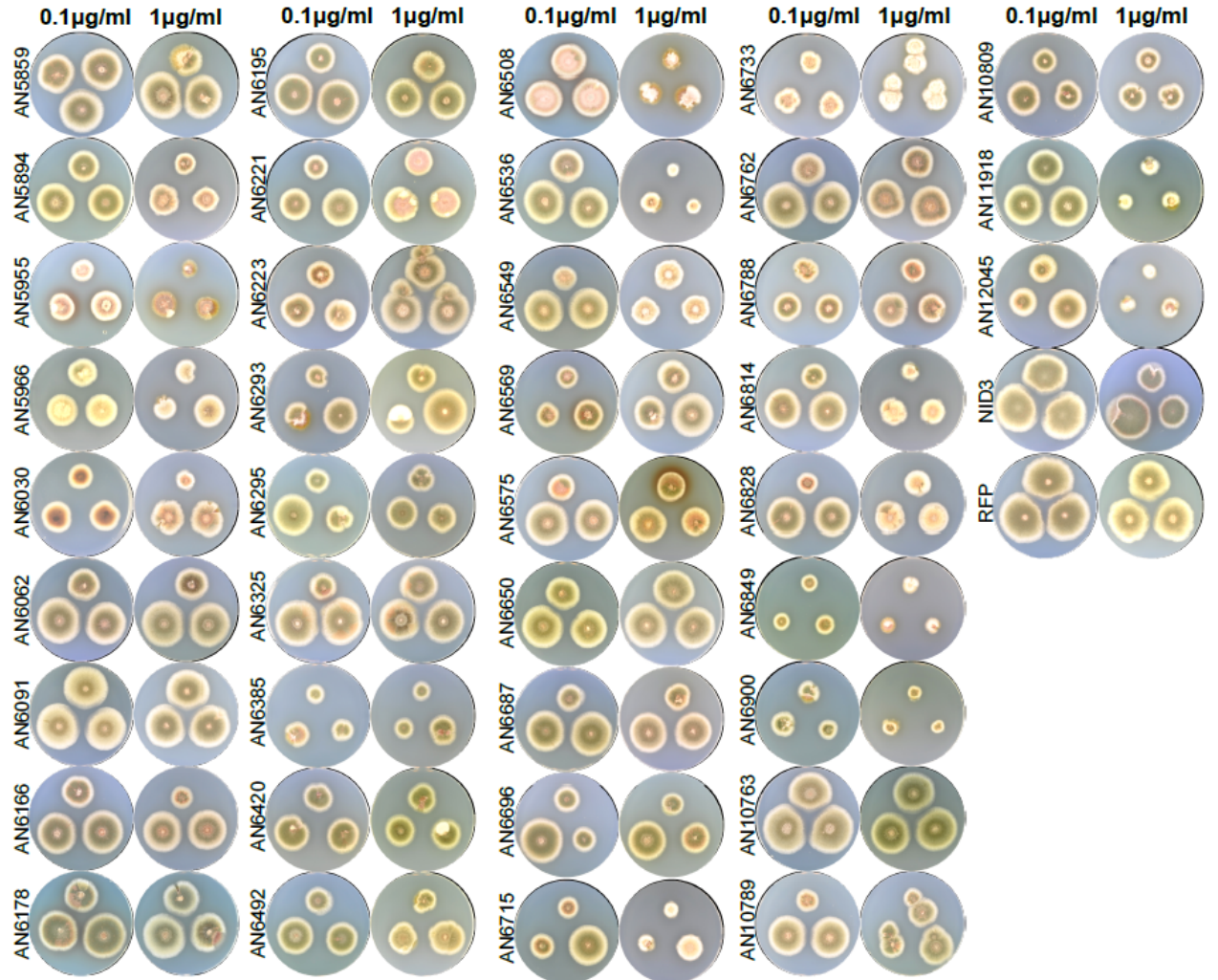


Figure 6.5: Macromorphologies of TF over-expression strains cultivated at two different concentrations of DOX. NID3 = reference with no insert, RFP = reference with *rfp* inserted.

Table 6.2: Average colony diameters (mm) of TF over-expression strains.

Strain	0.1 $\mu\text{g/ml}$ DOX	1.0 $\mu\text{g/ml}$ DOX	Comments
NID3	45 \pm 2	34 \pm 10	
RFP	50 \pm 1	40 \pm 2	
AN5859	40 \pm 3	38 \pm 6	
AN5894	35 \pm 3	24 \pm 5	
AN5955	25 \pm 4	21 \pm 4	Poor conidiation, production of red pigment
AN5966	30 \pm 1	24 \pm 7	Poor conidiation
AN6030	26 \pm 3	28 \pm 5	Poor conidiation, production of red pigment
AN6062	39 \pm 7	39 \pm 4	
AN6091	42 \pm 1	41 \pm 0	
AN6166	38 \pm 5	33 \pm 6	
AN6178	39 \pm 7	39 \pm 4	
AN6195 (<i>creA</i>)	37 \pm 8	35 \pm 2	
AN6221 (<i>areB</i>)	29 \pm 6	30 \pm 3	
AN6223	30 \pm 2	42 \pm 2	
AN6293 (<i>acuM</i>)	29 \pm 7	33 \pm 10	
AN6295	31 \pm 10	30 \pm 2	
AN6325	38 \pm 12	40 \pm 3	
AN6385	21 \pm 3	22 \pm 7	Poor growth
AN6420	33 \pm 2	30 \pm 0	
AN6492 (<i>hapE</i>)	32 \pm 4	28 \pm 4	
AN6508 (<i>gsk3</i>)	35 \pm 1	22 \pm 4	Poor conidiation
AN6536 (<i>hisB</i>)	36 \pm 4	16 \pm 4	
AN6549	34 \pm 10	27 \pm 1	
AN6569	25 \pm 4	36 \pm 6	
AN6575	34 \pm 11	32 \pm 9	Production of red pigment
AN6650 (<i>mscA</i>)	36 \pm 4	41 \pm 1	
AN6687	36 \pm 8	36 \pm 5	
AN6696	30 \pm 12	30 \pm 8	
AN6715	28 \pm 10	19 \pm 7	
AN6733	23 \pm 3	21 \pm 0	No conidiation, uneven colony periphery, poor growth
AN6762	39 \pm 6	37 \pm 4	
AN6788	26 \pm 1	28 \pm 3	

Continued on Next Page

Table 6.2 – Continued

Strain	0.1 $\mu\text{g/ml}$ DOX	1.0 $\mu\text{g/ml}$ DOX	Comments
AN6814	34 ± 9	24 ± 4	
AN6828	32 ± 10	30 ± 1	
AN6849	18 ± 2	15 ± 1	Poor growth
AN6900 (<i>tpiA</i>)	22 ± 2	15 ± 1	Poor growth
AN10763 (<i>adaB</i>)	45 ± 3	41 ± 1	Seems unaffected
AN10789	33 ± 11	29 ± 6	
AN10809	29 ± 5	26 ± 1	
AN11918/AN6430	38 ± 2	18 ± 2	
AN12045	30 ± 7	16 ± 3	

6.2.3 Metabolite profile analysis reveals production of novel metabolites

Metabolite profiles were analyzed by UHPLC-DAD-TOFMS, and the chromatograms were visually inspected for drastic changes in the metabolite profiles. All chromatograms (ESI⁺, ESI⁻, DAD) are shown in Appendix B. Based on the overall metabolite profiles I have selected two phosphatase strains (AN3793 (*ppzA*), AN8820 (*cnaA*)) and four TF strains (AN5955, AN6221 (*areB*)), AN6420, AN6733) to examine further. BPC (ESI⁺, ESI⁻) and DAD chromatograms of these six strains are shown in Figures 6.6, 6.7, and 6.8, respectively.

Examination of the chromatograms revealed that these six strains did indeed have altered metabolite profiles. Normally, only austinol, dehydroaustinol, and sterigmatocystin are produced in reasonable amounts on MM (see Figure 6.1), but apparently over-expression of both phosphatases and TFs can induce expression of other known metabolites on MM. Over-expression of AN3793 (phosphatase) and AN6733 (TF) induces expression of the antibiotic desferritriacetylfulsigen (Figure 6.9), which is not normally observed on standard media. AN8820 (*cnaA*, phosphatase) induces expression of arugosins and shamixanthones, which are products of the monodictyphenone pathway, and AN5955 (TF) induces expression of violaceols, which are derived from orsellinic acid (Chapter 5). Over-expression of AN6420 results in production of asperugin B, which is a polyketide that was only recently discovered and is also believed to origin from orsellinic acid [Klejnstrup, 2012]. It is noteworthy that both AN3793, AN8820 (*cnaA*), and AN6221 produce high amounts of sterigmatocystin, and in addition they produce high amounts of averantin, averufin, and versicolorin B, which are intermediates in the pathway towards sterigmatocystin in *A. nidulans*, and are only detected in trace amounts in a wild-type.

Interestingly, some of these strains also produce novel metabolites. Dereplication of the six metabolite profiles identified four novel metabolites, merely by manual inspection of the metabolite profiles. Three of the metabolites, designated Unknown 6, Unknown 22, and Unknown 26, have been detected in *A. nidulans*

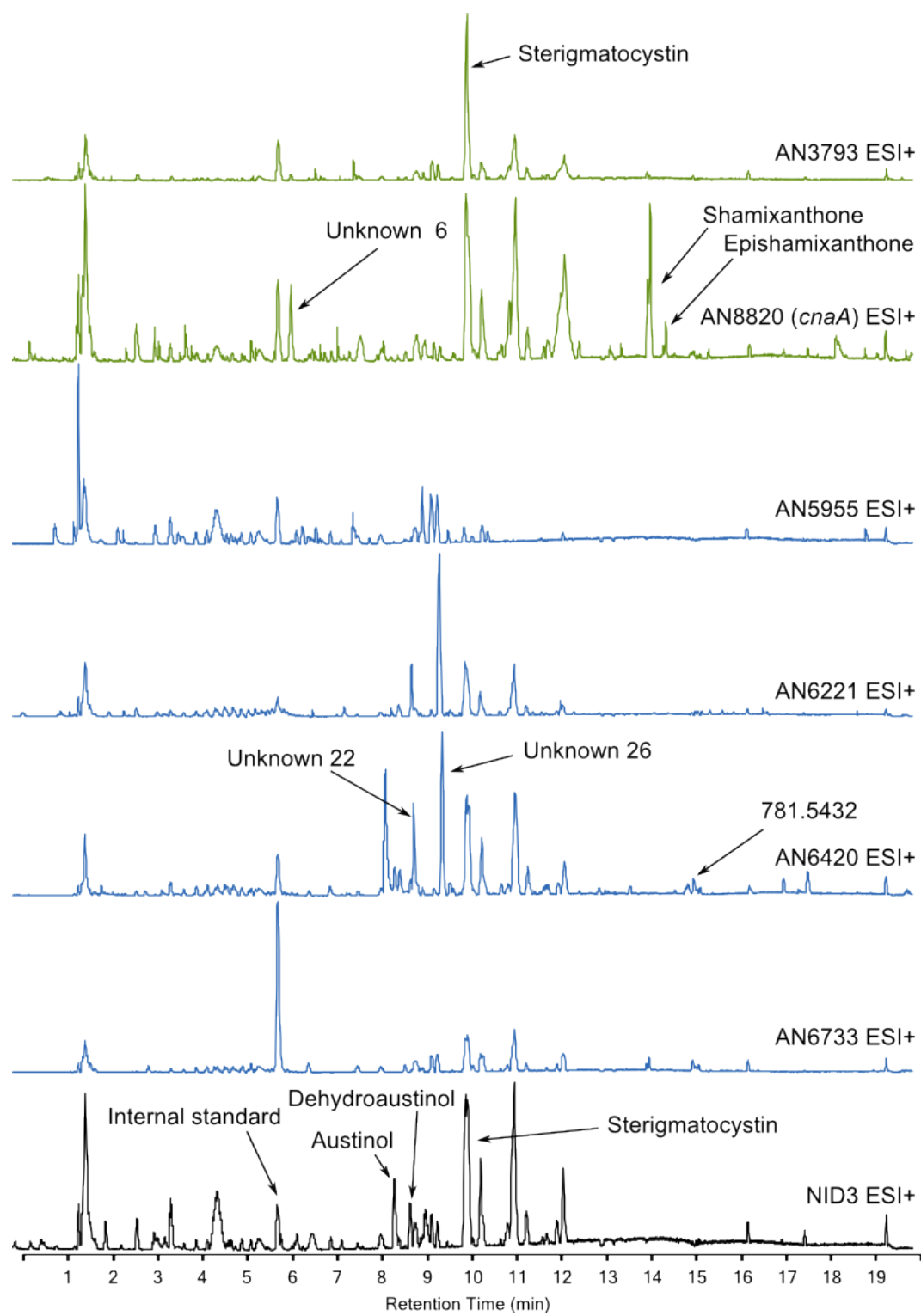


Figure 6.6: ESI+ chromatogram of six selected strains cultivated on 1.0 $\mu\text{g}/\text{ml}$ DOX with a NID3 reference.

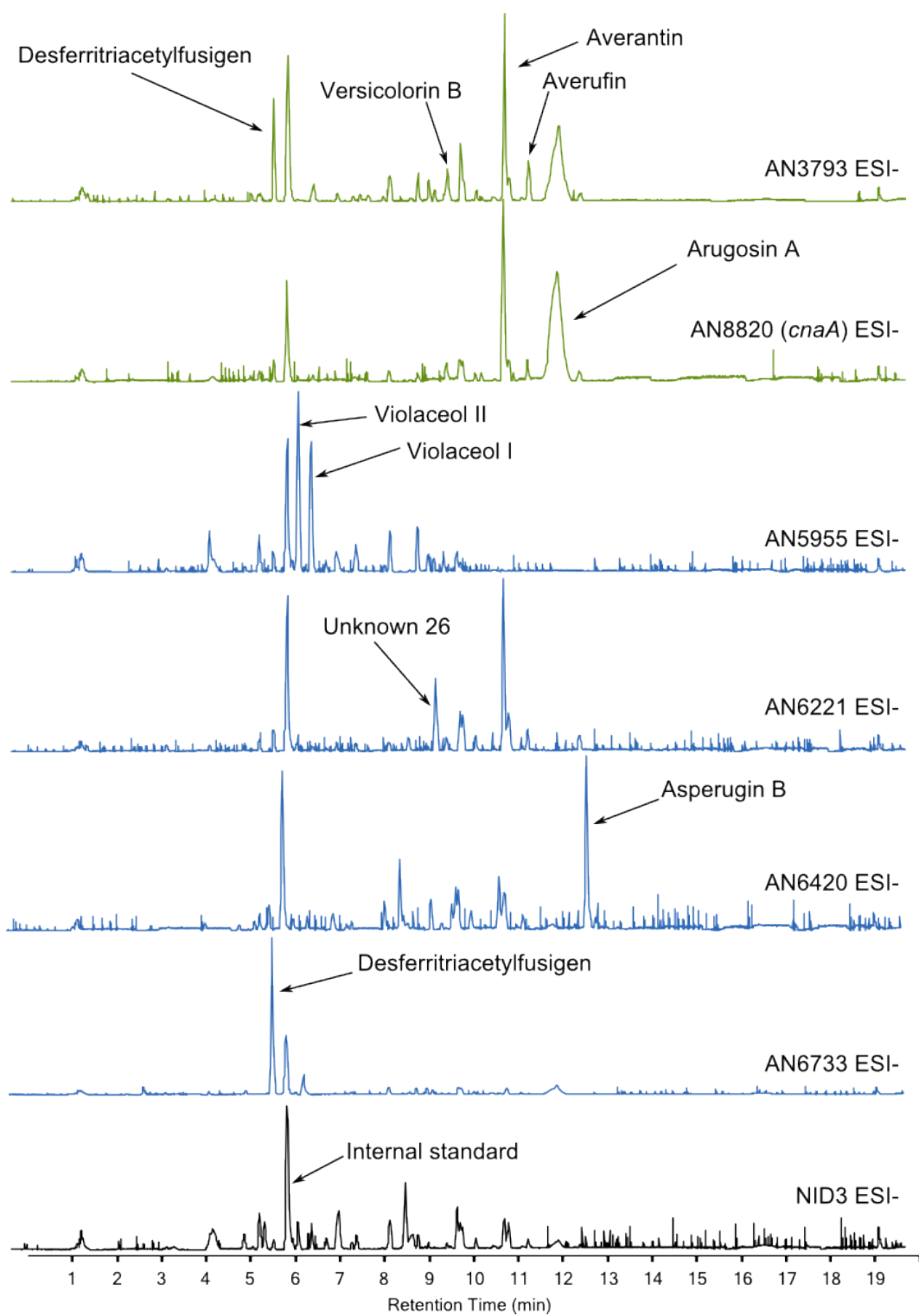


Figure 6.7: ESI- chromatogram of six selected strains cultivated on 1.0 $\mu\text{g}/\text{ml}$ DOX with a NID3 reference.

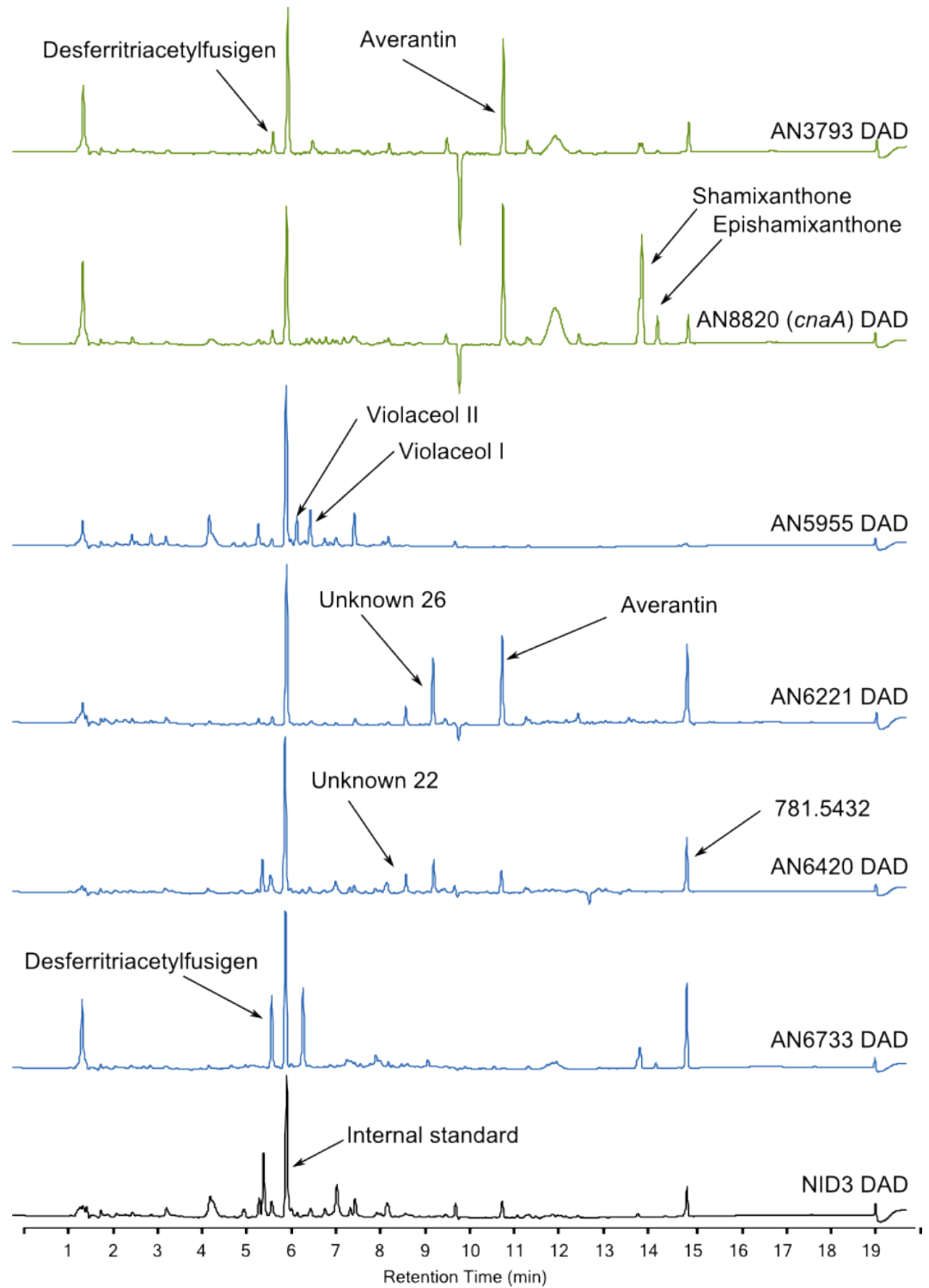


Figure 6.8: DAD chromatogram of six selected strains cultivated on 1.0 $\mu\text{g}/\text{ml}$ DOX with a NID3 reference.

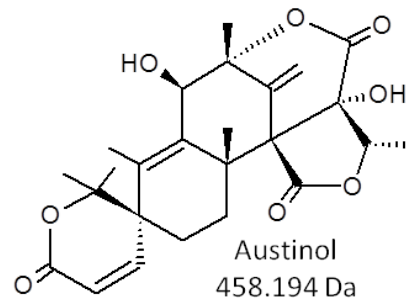
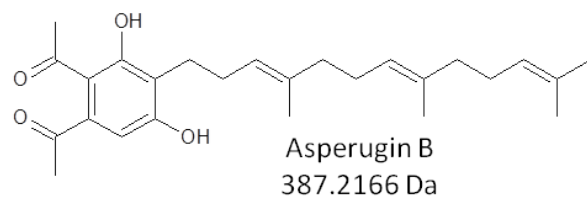
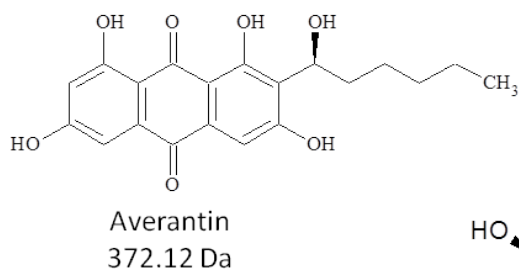
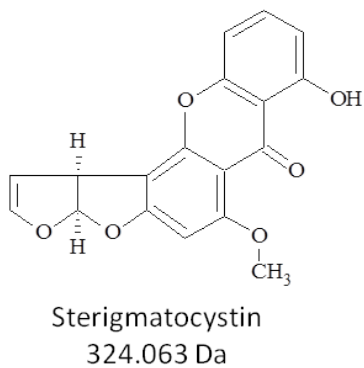
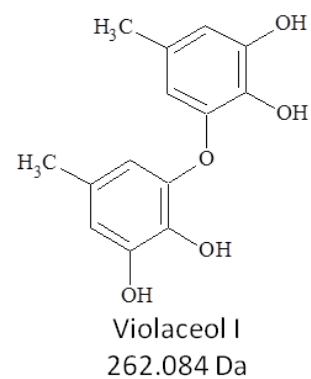
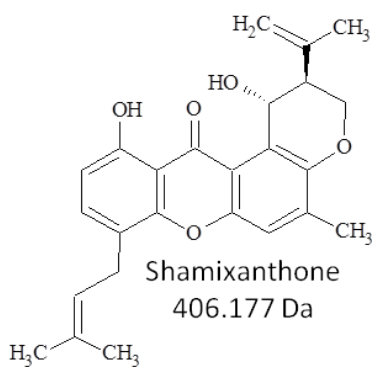
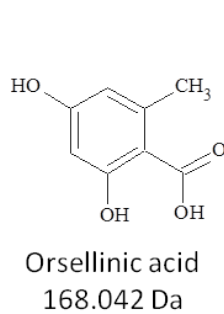
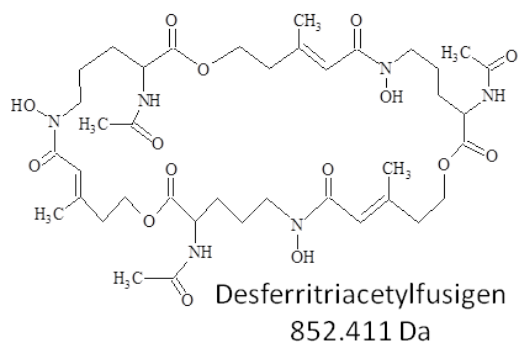
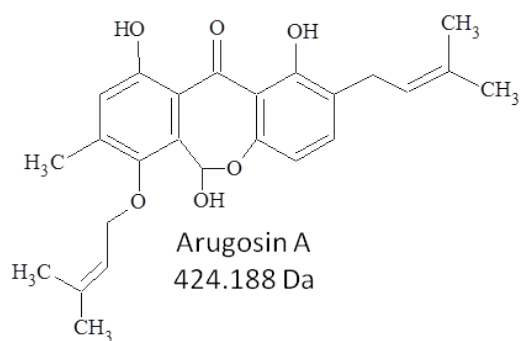


Figure 6.9: Selection of known *A. nidulans* compounds.

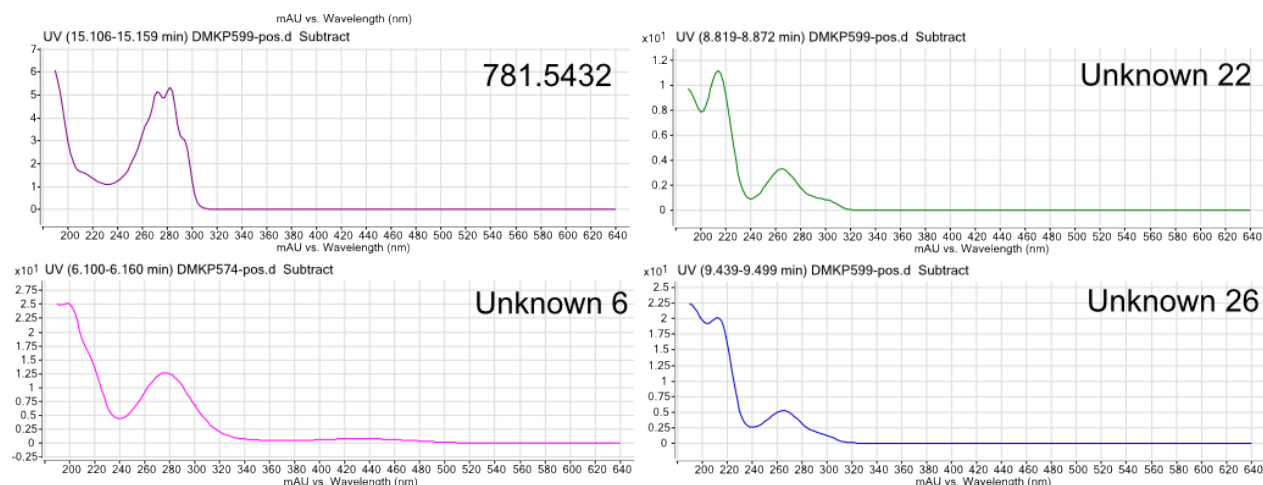


Figure 6.10: UV spectra of four compounds which were not identified by dereplication.

before (personal communication, Kristian F. Nielsen), but their structures are unknown. The compound designated 781.5432 indicates the monoisotopic mass of a compound, which has not been detected before. Inspection of the UV spectra of the unknown compounds (Figure 6.10) revealed that Unknown 6, Unknown 22, and Unknown 26 have similar UV spectra, why they may be analogues of the same compound. Based on the modest ability to absorb light, the compounds are most likely not polyketides, as these tend to be highly conjugated systems. Rather, these three compounds may be alkaloids/NRPs, and perhaps they are novel compounds with a potential use as pharmaceuticals.

Dereplication of 781.5432 did not identify any known metabolites why this compound may be a novel compound. Intriguingly, the UV spectrum of this compound is characteristic for ergosterol, a component of fungal cell walls, but the monoisotopic mass is almost twice that of ergosterol. Very recently, Miao and co-workers reported a steroid-xanthone heterodimer (asperversin A) from *A. versicolor*, which is the first of its kind [Miao et al., 2012]. Asperversin A is a heterodimer of $5\alpha,8\alpha$ -epidioxyergosta-6,22-dien-3 β -ol and 5-methoxysterigmatocystin, joined by an ether bond (Figure 6.11). Although this compound only differs from 781.5432 by 1 Dalton, the UV spectrum dismisses this compound as a candidate, as asperversin A has far greater UV absorbance due to the xanthone monomer (personal communication, Jens C. Frisvad). In fact, based on the hypothesis that the UV spectrum of ergosterol or derivatives will not be greatly influenced by formation of a heterodimer, the counterpart of such a compound would necessarily have very poor UV absorbance e.g. a NRP.

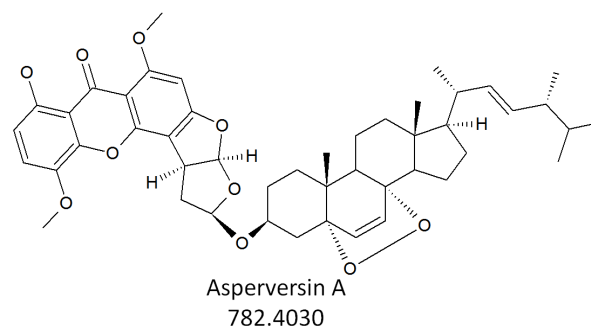


Figure 6.11: Structure of the recently reported aspersversin A; a heterodimer of 5 α ,8 α -epidioxyergosta-6,22-dien-3 β -ol and 5-methoxysterigmatocystin.

None of the six examined strains seem to be part of a secondary metabolite gene cluster i.e. there are no putative PKSs, NRPSs, terpene synthases, prenyltransferases, or cytochrome P450s up- and downstream of these genes (± 10 genes). There is a putative PKS located 11 genes from AN6420 (AN6431), however, the AN6420 over-expression strain seem to only upregulate NRP and the 781.5432 compounds, which is also not likely a polyketide.

The two phosphatases are both characterized proteins. Protein phosphatase Z (AN3793, *ppzA*) is involved in oxidative stress response and belongs to a group of proteins that are specific to fungi [Leiter et al., 2012]. Over-expression of *ppzA* upregulates production of desferritriacetylfusigen and sterigmatocystin, the latter in a manner that may exceed the capacity of tailoring enzymes, as the known intermediates averantin and averufin accumulate in high amounts. The link between these metabolites and Protein phosphatase Z is not obvious.

cnaA (AN8820) encodes a calmodulin-dependent protein phosphatase catalytic subunit, which is an essential protein involved in cell cycle regulation [Rasmussen et al., 1994, Nanthakumar et al., 1996, Son and Osmani, 2009, Hernández-Ortiz and Espeso, 2013]. Over-expression of *cnaA* results in up-regulation of shamixanthone production, however, a possible link between shamixanthenes and cell cycle regulation is also not obvious.

areB (AN6221) encodes a characterized transcription factor AreB, which is involved in nitrogen metabolism [Conlon et al., 2001, Wong et al., 2009]. It has previously been shown that AreB is pleiotropic and that over-expression of *areB* results in severe growth inhibition and aberrant cell morphology [Wong et al., 2009]. Wong and co-workers over-expressed the *areB* gene under control of a xylose-inducible promoter, and integrated the construct into the *yA* locus. Interestingly, in the *areB* over-expression strain, there is a substantial up-regulation of Unknown 26, which appears to be a NRP, which is a product of the nitrogen metabolism. At this point it is impossible to propose the exact connection, and the fact that *areB* encodes at least three different proteins [Conlon et al., 2001] illustrates the complexity of this regulation.

6.3 Concluding remarks

In this study I successfully applied the HTP setup that I have developed, for over-expression of *A. nidulans* phosphatases and transcription factors. The HTP setup proved very efficient and strain construction was very fast and only little manual work was required. Due to problems with protoplasting and initial PCR amplification not all intended strain were constructed. Also, eight of the constructed strains were not analyzed by UHPLC-DAD-TOFMS due to technical problems with the TOFMS. However, 52 strains were obtained and several of them did show an altered metabolite profile, resulting in either overall down-regulation of secondary metabolite production, or up-regulation of single or a few metabolites which are normally not produced under the specific circumstances. Furthermore, examination of six selected strain identified four possible novel compounds. As this experiment was the last one I carried out during my PhD, I did not have sufficient time for construction of the missing strains or in-depth analysis of all analytical data.

At this point in time the aspects of regulation is apparently too complex for researchers to elucidate the remaining parts of fungal secondary metabolism, why I predict that random HTP approaches can be very helpful in uncovering some of the mechanisms involved with regulation of secondary metabolism, and this study confirms that regulation by transcriptional control and/or protein phosphorylation is part of this regulation. Additionally, these libraries may be of use to the entire fungal research community for studies within cell cycle regulation, cell differentiation etc.

6.4 Experimental procedures

6.4.1 Fungal strains

The *A. nidulans* strain NID1 (IBT 29539) (*argB2 pyrG89 veA1 nkuAΔ*) [Nielsen et al., 2008] was used as a reference and for insertion of all genes. The recipient strain (NIDX) for background-free integration was constructed by integration of a *wA::pyrG* cassette into IS1, as described in Chapter 3.

Table 6.3: Fungal strains used and constructed in this study. *hph** contains a silent point mutation eliminating an AsiSI restriction site. *ori^R** contains a point mutation eliminating a Nb.BtsI nicking site.

Strain	Genotype	Source
NIDX	<i>argB2 pyrG89 veA1 nkuAΔ</i> IS1::P <i>gpdA</i> _{0.5kb} :: <i>wA::TtrpC::pyrG</i>	Chapter 3
RFP	<i>argB2 pyrG89 veA1 nkuAΔ</i> IS1::P <i>O7::rfp::TtrpC::argB</i>	This study
GFP	<i>argB2 pyrG89 veA1 nkuAΔ</i> IS1::P <i>O7::gfp::TtrpC::argB</i>	This study
Phosphatase strains		
AN0129	<i>argB2 pyrG89 veA1 nkuAΔ</i> IS1::P <i>O7::AN0129::TtrpC::argB</i>	This study
AN0164	<i>argB2 pyrG89 veA1 nkuAΔ</i> IS1::P <i>O7::AN0164::TtrpC::argB</i>	This study
AN0914	<i>argB2 pyrG89 veA1 nkuAΔ</i> IS1::P <i>O7::AN0914::TtrpC::argB</i>	This study
AN3793	<i>argB2 pyrG89 veA1 nkuAΔ</i> IS1::P <i>O7::AN3793::TtrpC::argB</i>	This study
AN3810	<i>argB2 pyrG89 veA1 nkuAΔ</i> IS1::P <i>O7::AN3810::TtrpC::argB</i>	This study
AN3941 (nimT)	<i>argB2 pyrG89 veA1 nkuAΔ</i> IS1::P <i>O7::AN3941::TtrpC::argB</i>	This study
AN4426	<i>argB2 pyrG89 veA1 nkuAΔ</i> IS1::P <i>O7::AN4426::TtrpC::argB</i>	This study
AN5057	<i>argB2 pyrG89 veA1 nkuAΔ</i> IS1::P <i>O7::AN5057::TtrpC::argB</i>	This study
AN6892	<i>argB2 pyrG89 veA1 nkuAΔ</i> IS1::P <i>O7::AN6892::TtrpC::argB</i>	This study
AN8820 (cnaA)	<i>argB2 pyrG89 veA1 nkuAΔ</i> IS1::P <i>O7::AN8820::TtrpC::argB</i>	This study
AN10138	<i>argB2 pyrG89 veA1 nkuAΔ</i> IS1::P <i>O7::AN10138::TtrpC::argB</i>	This study
AN10281	<i>argB2 pyrG89 veA1 nkuAΔ</i> IS1::P <i>O7::AN10281::TtrpC::argB</i>	This study
AN10570	<i>argB2 pyrG89 veA1 nkuAΔ</i> IS1::P <i>O7::AN10570::TtrpC::argB</i>	This study
TF strains		
AN5859	<i>argB2 pyrG89 veA1 nkuAΔ</i> IS1::P <i>O7::AN5859::TtrpC::argB</i>	This study
AN5894	<i>argB2 pyrG89 veA1 nkuAΔ</i> IS1::P <i>O7::AN5894::TtrpC::argB</i>	This study
AN5955	<i>argB2 pyrG89 veA1 nkuAΔ</i> IS1::P <i>O7::AN5955::TtrpC::argB</i>	This study
AN5966	<i>argB2 pyrG89 veA1 nkuAΔ</i> IS1::P <i>O7::AN5966::TtrpC::argB</i>	This study
AN6030	<i>argB2 pyrG89 veA1 nkuAΔ</i> IS1::P <i>O7::AN6030::TtrpC::argB</i>	This study
AN6062	<i>argB2 pyrG89 veA1 nkuAΔ</i> IS1::P <i>O7::AN6062::TtrpC::argB</i>	This study

Continued on Next Page

Table 6.3 – Continued

Strain	Genotype	Source
AN6091	<i>argB2 pyrG89 veA1 nkuAΔ</i> IS1::P O7::AN6091::T <i>trpC</i> :: <i>argB</i>	This study
AN6166	<i>argB2 pyrG89 veA1 nkuAΔ</i> IS1::P O7::AN6166::T <i>trpC</i> :: <i>argB</i>	This study
AN6178	<i>argB2 pyrG89 veA1 nkuAΔ</i> IS1::P O7::AN6178::T <i>trpC</i> :: <i>argB</i>	This study
AN6195	<i>argB2 pyrG89 veA1 nkuAΔ</i> IS1::P O7::AN6195::T <i>trpC</i> :: <i>argB</i>	This study
AN6221	<i>argB2 pyrG89 veA1 nkuAΔ</i> IS1::P O7::AN6221::T <i>trpC</i> :: <i>argB</i>	This study
AN6223	<i>argB2 pyrG89 veA1 nkuAΔ</i> IS1::P O7::AN6223::T <i>trpC</i> :: <i>argB</i>	This study
AN6293	<i>argB2 pyrG89 veA1 nkuAΔ</i> IS1::P O7::AN6293::T <i>trpC</i> :: <i>argB</i>	This study
AN6295	<i>argB2 pyrG89 veA1 nkuAΔ</i> IS1::P O7::AN6295::T <i>trpC</i> :: <i>argB</i>	This study
AN6325	<i>argB2 pyrG89 veA1 nkuAΔ</i> IS1::P O7::AN6325::T <i>trpC</i> :: <i>argB</i>	This study
AN6385	<i>argB2 pyrG89 veA1 nkuAΔ</i> IS1::P O7::AN6385::T <i>trpC</i> :: <i>argB</i>	This study
AN6420	<i>argB2 pyrG89 veA1 nkuAΔ</i> IS1::P O7::AN6420::T <i>trpC</i> :: <i>argB</i>	This study
AN6492	<i>argB2 pyrG89 veA1 nkuAΔ</i> IS1::P O7::AN6492::T <i>trpC</i> :: <i>argB</i>	This study
AN6508	<i>argB2 pyrG89 veA1 nkuAΔ</i> IS1::P O7::AN6508::T <i>trpC</i> :: <i>argB</i>	This study
AN6536	<i>argB2 pyrG89 veA1 nkuAΔ</i> IS1::P O7::AN6536::T <i>trpC</i> :: <i>argB</i>	This study
AN6549	<i>argB2 pyrG89 veA1 nkuAΔ</i> IS1::P O7::AN6549::T <i>trpC</i> :: <i>argB</i>	This study
AN6569	<i>argB2 pyrG89 veA1 nkuAΔ</i> IS1::P O7::AN6569::T <i>trpC</i> :: <i>argB</i>	This study
AN6575	<i>argB2 pyrG89 veA1 nkuAΔ</i> IS1::P O7::AN6575::T <i>trpC</i> :: <i>argB</i>	This study
AN6650	<i>argB2 pyrG89 veA1 nkuAΔ</i> IS1::P O7::AN6650::T <i>trpC</i> :: <i>argB</i>	This study
AN6687	<i>argB2 pyrG89 veA1 nkuAΔ</i> IS1::P O7::AN6687::T <i>trpC</i> :: <i>argB</i>	This study
AN6696	<i>argB2 pyrG89 veA1 nkuAΔ</i> IS1::P O7::AN6696::T <i>trpC</i> :: <i>argB</i>	This study
AN6715	<i>argB2 pyrG89 veA1 nkuAΔ</i> IS1::P O7::AN6715::T <i>trpC</i> :: <i>argB</i>	This study
AN6733	<i>argB2 pyrG89 veA1 nkuAΔ</i> IS1::P O7::AN6733::T <i>trpC</i> :: <i>argB</i>	This study
AN6762	<i>argB2 pyrG89 veA1 nkuAΔ</i> IS1::P O7::AN6762::T <i>trpC</i> :: <i>argB</i>	This study
AN6788	<i>argB2 pyrG89 veA1 nkuAΔ</i> IS1::P O7::AN6788::T <i>trpC</i> :: <i>argB</i>	This study
AN6814	<i>argB2 pyrG89 veA1 nkuAΔ</i> IS1::P O7::AN6814::T <i>trpC</i> :: <i>argB</i>	This study
AN6828	<i>argB2 pyrG89 veA1 nkuAΔ</i> IS1::P O7::AN6828::T <i>trpC</i> :: <i>argB</i>	This study
AN6849	<i>argB2 pyrG89 veA1 nkuAΔ</i> IS1::P O7::AN6849::T <i>trpC</i> :: <i>argB</i>	This study
AN6900	<i>argB2 pyrG89 veA1 nkuAΔ</i> IS1::P O7::AN6900::T <i>trpC</i> :: <i>argB</i>	This study
AN10763	<i>argB2 pyrG89 veA1 nkuAΔ</i> IS1::P O7::AN10763::T <i>trpC</i> :: <i>argB</i>	This study
AN10789	<i>argB2 pyrG89 veA1 nkuAΔ</i> IS1::P O7::AN10789::T <i>trpC</i> :: <i>argB</i>	This study
AN10809	<i>argB2 pyrG89 veA1 nkuAΔ</i> IS1::P O7::AN10809::T <i>trpC</i> :: <i>argB</i>	This study
AN11918	<i>argB2 pyrG89 veA1 nkuAΔ</i> IS1::P O7::AN11918::T <i>trpC</i> :: <i>argB</i>	This study
AN12045	<i>argB2 pyrG89 veA1 nkuAΔ</i> IS1::P O7::AN12045::T <i>trpC</i> :: <i>argB</i>	This study

6.4.2 Media

Minimal medium for *A. nidulans* was made as described in [Cove, 1966] but with 1% glucose, 10 mM NaNO₃ and 2% agar. Minimal medium for *A. niger* was prepared as described in [Chiang et al., 2011]. YES, MEA and CYA media were prepared as described by Frisvad and Samson [Samson et al.]. When necessary, media were supplemented with 4 mM L-arginine, 10 mM uridine, 10 mM uracil. Furthermore, media were supplemented with 0.1-20 μ g/ml doxycycline hyclate (doxycycline hydrochloride hemiethanolate hemihydrate) (Sigma-Aldrich).

6.4.3 PCR, cloning and transformation

All primers are listed in Table 3.3.3. All PCR reactions were carried out using the *PfuX7* polymerase [Nørholm, 2010], and either HF or Phire PCR buffer (Finnzymes/Thermo Scientific). For phosphatases, primer pairs named ANxxxx-Fw and ANxxxx-Rv were used for amplification of ORFs, and for transcription factors primer pairs named F-xxxx and R-xxxx were used for amplification of ORFs from *A. nidulans* genomic DNA (gDNA). All primers are listed in Table 6.4.3. The PCR program was: 98°C for 1 min, 98°C for 15 s, 60°C for 30 s, 72°C for 30s/kb, and 72°C for 5 min. Normally, I performed 35 cycles of amplification. Diagnostic (conidial) PCR was performed as a regular PCR reaction, except initial denaturation step was 20 min. PCR reactions were mixed in a Hamilton STAR liquid handler in a total volume of 50 μ l.

Length and quality of the PCR fragments was assessed via capillary electrophoresis on a Lab-Chip GX (Caliper LifeSciences) and compared to the calculated lengths of PCR products, which was calculated using the novel program VirtualPCR, developed for this study by Ali Altintas, Center for Biological Sequence Analysis, DTU.

Unpurified PCR fragments were assembled via uracil-excision fusion [Geu-Flores et al., 2007] into the compatible vector pDH124. The background-free cloning vector pDH124 was constructed by insertion of *ccdB::camR* into the UEC of pDCA-Tet (kindly provided by Diana C. Anyaogu) as described in Section 3.1.5. Cloning reactions were carried out in microtiter PCR plates, and one reaction (10 μ l) contained 1 μ l USER enzyme, 0.5 μ l New England Biolabs buffer 4 (NEB4), 0.5 μ l 10x bovine serum albumin (BSA), 2 μ l unpurified PCR fragment in HF buffer, and 6 μ l MQ water. Correct clones were verified by restriction digestion. Plasmids were linearized by digestion with NotI, heat inactivated, and transformed directly into *A. nidulans*. Protoplasting and gene-targeting procedures were performed as described previously [Johnstone et al., 1985, Nielsen et al., 2006]. All plasmids are listed in Table 6.4.3.

Table 6.4: Primers used in this study. All primers are shown in 5'-3' direction.

Primer name	Sequence
mRFP-ORF-Fw	AGAGCGAUATGGCCTCCTCCGAGGAC
mRFP-ORF-Rv	TCTGCGAUTTAGGCGCCGGTGGAGT
GFP-ORF-Fw	AGAGCGAUATGGTGAGCAAGGGCGAG
GFP-ORF-Rv	TCTGCGAUTTACTTGTACAGCTCGTCCATGC
GFP-fw1	AGCTGGUGCAGGCGCTGGAGCC
Stag-rv1	TCTGCGAUTTAAGCACCGCTGTCCATGT
AN0103-fw	AGAGCGAUATGAGTGATCTGGATAGGTACCGT
AN0129-fw	AGAGCGAUATGGCAACCGTTGTGGTG
AN0164-fw	AGAGCGAUATGCCTGGTTTACCAGGTACGT
AN0410-fw	AGAGCGAUATGGCCGATCAGGACGTG
AN0504-fw	AGAGCGAUATGACTGATAGCAAGGTCCCTC
AN0914-fw	AGAGCGAUATGACTCTGGCTCCGTGCGAA
AN1343-fw	AGAGCGAUATGAACTCGCTGAATATCCTATCG
AN1358-fw	AGAGCGAUATGGGCCAGACTCTGTGTCAGAG
AN1467-fw	AGAGCGAUATGATCTTATCGGTCAGGCCC
AN2472-fw	AGAGCGAUATGGACAGATCGAATACAGTGGA
AN2902-fw	AGAGCGAUATGCTTCTTCGTTTACCGCCA
AN3793-fw	AGAGCGAUATGGGTCAGTCTCATTCOAAGG
AN3810-fw	AGAGCGAUATGGCCTACGACCCGCGA
AN3941-fw	AGAGCGAUATGGAACATTCCTCGCCTTT
AN4419-fw	AGAGCGAUATGGCTATGAACAAGGTTCCC
AN4426-fw	AGAGCGAUATGACTTACCCTTTGACGAAGAAA
AN4544-fw	AGAGCGAUATGCCCGTTCAAACCAGC
AN4896-fw	AGAGCGAUATGGCGCATCCCACCATT
AN5057-fw	AGAGCGAUATGGCCGGATCTACTGCTG
AN5722-fw	AGAGCGAUATGCGTCGTGCAGCCCT
AN6391-fw	AGAGCGAUATGGATAACAATATGGAGATCGATG
AN6892-fw	AGAGCGAUATGTTTCAGTGGCTCCTCGAG
AN6982-fw	AGAGCGAUATGTCTGCGATGACAGGCC
AN8820-fw	AGAGCGAUATGGATCGGAATCTAGCGC
AN10077-fw	AGAGCGAUATGTTTCGGCAAGAGTCCAGAAC
AN10138-fw	AGAGCGAUATGGAACCTACAGCTTCAGGCT

Continued on Next Page

Table 6.4 – Continued

Primer name	Sequence
AN10281-fw	AGAGCGAUATGGCCTCCCAAGATGTTGA
AN10570-fw	AGAGCGAUATGTACTGGGTATACTTCGTGATGG
AN0103-rv	ACCAGCUCCTTAGACGGAGAAGATTTGAGGG
AN0129-rv	ACCAGCUCCTTATCTCGAATAAGGCTTGTTTCAT
AN0164-rv	ACCAGCUCCTCACAGAAAGTAGTCGATCACAGGA
AN0410-rv	ACCAGCUCCTTACTTCTTTTGCTTTCGCGGA
AN0504-rv	ACCAGCUCCTTACAAAAAGTACTCGCTGCG
AN0914-rv	ACCAGCUCCTCATCTCCGTGACCTGCG
AN1343-rv	ACCAGCUCCTCATAGAGCGTCTGCTTCTCCT
AN1358-rv	ACCAGCUCCTTACGAAGAGGCGGGTTTCT
AN1467-rv	ACCAGCUCCTACTCCTCCACAGCAATTATAACCA
AN2472-rv	ACCAGCUCCTTAGGCCAACCTCGCCTC
AN2902-rv	ACCAGCUCCTCAGCCTTTATCCTCGCCAA
AN3793-rv	ACCAGCUCCTTAAAAGCTTTGGGCCGAAAC
AN3810-rv	ACCAGCUCCTTACAACCAAGCCAAAGTCCATA
AN3941-rv	ACCAGCUCCTTAGTATGAAAGCATGCGACG
AN4419-rv	ACCAGCUCCTAATGATCAAAGCTGTACAAGACGTAC
AN4426-rv	ACCAGCUCCTCATGAAGCGACGCGAATC
AN4544-rv	ACCAGCUCCTCACAGAAACTCATCAATGCTTCT
AN4896-rv	ACCAGCUCCTCAATCAGTATCAGAGTGCTCAGG
AN5057-rv	ACCAGCUCCTCATGCTTGCGCCTTTACAC
AN5722-rv	ACCAGCUCCTTAAAGCTTAGCTTTGATGTTATCTTC
AN6391-rv	ACCAGCUCCTCACAGGAAGTAGTCTGGGGTAC
AN6892-rv	ACCAGCUCCTCATGTACTGTCGCTCGAACC
AN6982-rv	ACCAGCUCCTTATCCCGGGCTCTCCCT
AN8820-rv	ACCAGCUCCTTACCGTCTCGTTGTGCGACG
AN10077-rv	ACCAGCUCCTCATAGTGTGATGTCGAGAACCAG
AN10138-rv	ACCAGCUCCTTATGCCAGCGATCCTGA
AN10281-rv	ACCAGCUCCTCAACGAATTTTGCGCATAG
AN10570-rv	ACCAGCUCCTAATTCTGATTCTTCTTCTCCAAGTACT
F-5836	AGAGCGAUATGGCCAGCATGAATCAACCT
F-5859	AGAGCGAUATGGCTCCATCTTCTCGTCCT
F-5894	AGAGCGAUCCTCGAACCGCTGCCG

Continued on Next Page

Table 6.4 – Continued

Primer name	Sequence
F-5929	AGAGCGAUATGGACAACAGGCCTGCTTC
F-5955	AGAGCGAUATGGCCTCCAACAGCAAAG
F-5966	AGAGCGAUATGGATCCTAGAAACCATCCCT
F-6010	AGAGCGAUTCTTTCTCTCTCCATCCTCCATC
F-6030	AGAGCGAUATGGACGGATCACATAAGGAGA
F-6062	AGAGCGAUATGCCCCAGTTCAACTGCC
F-6091	AGAGCGAUCCTTCCCCTCTCGTGCGAG
F-6166	TCTGCGAUGAGATGATTTACATATTTTCGATT
F-6173	AGAGCGAUATGAACCCTTCTGACTTGGCT
F-6178	AGAGCGAUATGTCGCCCCCGACCTTA
F-6195	AGAGCGAUATTCCCTCGACTTCATACATCC
F-6221	AGAGCGAUATGGCTAGCACGACCTACG
F-6223	AGAGCGAUATGGCCGACATCCTCACC
F-6293	AGAGCGAUATGACGGAGAAAAACACTACGCA
F-6295	AGAGCGAUATGAGCATCAATCCCCCAAAT
F-6325	AGAGCGAUATGACGGACACTCGTCCAGTC
F-6385	AGAGCGAUATGAGCATCTTAGGTATTATAACA
F-6396	AGAGCGAUATGAGTGACACCTTATCCACC
F-6420	AGAGCGAUATGAGCAAGCCCCTTACA
F-6492	AGAGCGAUCAGATCCTGCTTGTATTGCGT
F-6503	AGAGCGAUATGGAGACCGCAGAGCCT
F-6508	AGAGCGAUCTTCTTCACGTTGTCCGC
F-6536	AGAGCGAUATGCCCCTTCCTGCCCCG
F-6549	AGAGCGAUATGGCAGGCACTCAGGAC
F-6566	AGAGCGAUATGGAGCAGCCTAATGAACCG
F-6569	AGAGCGAUATGATTGCCATCGGTCTTGAAG
F-6575	AGAGCGAUATGACAACGACGACGGCA
F-6650	AGAGCGAUTTCCTCATTGATTCACATTCTTCAGTTC
F-6687	AGAGCGAUATGTGTGCTTCACCGAATTTCG
F-6696	AGAGCGAUATGGAGTCGTCTCCCCCGA
F-6715	AGAGCGAUATGACTACCAGCAATCACCATCAA
F-6733	AGAGCGAUATGGCCTTTGCGGTGACG
F-6747	AGAGCGAUATGCCCGCTGAAGAGAGG

Continued on Next Page

Table 6.4 – Continued

Primer name	Sequence
F-6762	AGAGCGAUATGCCCCTTCGACGAGCA
F-6788	AGAGCGAUATGCCGCAGAAAGCCCAG
F-6814	AGAGCGAUATGTTTGAAGTGGAGACTCCGTC
F-6828	AGAGCGAUATGCTGTGTCACTTCTCGATG
F-6832	AGAGCGAUATGTTTCATCGCCTTCGAAGCTACC
F-6849	AGAGCGAUAATTCCATCAATCCTATCTATCCTGATCAT
F-6861	AGAGCGAUATGGCACCGGACATAGAG
F-6884	AGAGCGAUATGGATCTGGAGTGCCGG
F-6889	AGAGCGAUATGGCGACACACGACGG
F-6900	AGAGCGAUATGCCTCGCAAATTTTTCGTTGG
F-6960	AGAGCGAUATGACGAGTGACTCCTACGATCA
F-10763	AGAGCGAUTCCCTTTATCTGGGTATTATCTGC
F-10789	AGAGCGAUATGTCGTCTCTGCCCAGGA
F-10809	AGAGCGAUATGACAGAGCAAGACAACCTCCTATC
F-11918	AGAGCGAUATGATCTCCGCTCCGGGT
F-12045	AGAGCGAUATGTCAGCTAGTCCATCTCCGTC
R-5836	TCTGCGAUTTAACGACGAGCACTTATCAGACT
R-5859	TCTGCGAUTCATCCTGCAAACGGCG
R-5894	TCTGCGAUTTACTTGGATCCCCATCCG
R-5929	TCTGCGAUCTATTGCGGGTAGCTGGTCTG
R-5955	TCTGCGAUTCATTGTGCGAGCAACTGAAAC
R-5966	TCTGCGAUATAGTTGGACGCCGGATAAAC
R-6010	TCTGCGAUGTTAAATCGAGAGGAGAATTTCTAT
R-6030	TCTGCGAUTCACCAGATCTGATTCTCAGTAACTG
R-6062	TCTGCGAUTTAAGACAGGCATAGAACCTCTTTG
R-6091	TCTGCGAUGGGCAAGTGATCTACATTATGTGTT
R-6166	AGAGCGAUGCAACCTGCCATCTCACT
R-6173	TCTGCGAUCTACAGCGGGTTGAATGCC
R-6178	TCTGCGAUTCAGTTCACCTTCTTCTTTTTCG
R-6195	TCTGCGAUGTACAGTGTTCCGATTCAAGGAA
R-6221	TCTGCGAUCTAGATACTCTTAGGCGATTTGGCT
R-6223	TCTGCGAUCTAGTATGAGCTCCCCCATAGATA
R-6293	TCTGCGAUCTAGAAGAGGAGTTTTTGAATCGC

Continued on Next Page

Table 6.4 – Continued

Primer name	Sequence
R-6295	TCTGCGAUCATCAAGACATTATAATTTACTTAGTGGA
R-6325	TCTGCGAUGTAATAATTTAGATCAATGATTGATTC
R-6385	TCTGCGAUTCATGTCTGTTGTGAGTG
R-6396	TCTGCGAUTCAGTATAGATGGCTGGGAACG
R-6420	TCTGCGAUCTAGTCACTATCCAAGACAATAAC
R-6492	TCTGCGAUACGCAAAGAAGTAAATAATCAAAATGGGTA
R-6503	TCTGCGAUTTACTCGTCACCGCGCTT
R-6508	TCTGCGAUTCAGTCAAGGTGTGCCAT
R-6536	TCTGCGAUGAGTGGCGTGTTCATCGC
R-6549	TCTGCGAUCTAACTCGCCGTTTGTCTTTTTTC
R-6566	TCTGCGAUTTAAATTTGGTCTAATATCGCCGCATG
R-6569	TCTGCGAUCTAATCTCTCCATTGCACAAATACATCAT
R-6575	TCTGCGAUCAATGTGCAAATGACTAGGTTTGC
R-6650	TCTGCGAUTAGCGCAAATATAGACTCTATGTCTAG
R-6687	TCTGCGAUTTACTTCCCAGTCTTCGACTTC
R-6696	TCTGCGAUTCAATAACCCAGCATCGCAGAAAGC
R-6715	TCTGCGAUTCACCAGTCCTGGCCAC
R-6733	TCTGCGAUTTATTGAGAGACTTCCCTCAATTTATCAAT
R-6747	TCTGCGAUCTACTGCTTATCACTGTCCGG
R-6762	TCTGCGAUCTATAGGAAACGATACAGAAACTCCAGG
R-6788	TCTGCGAUCTAACCCGCAGTCCCACTTAC
R-6814	TCTGCGAUTTATCGCCGTGTCGAATGC
R-6828	TCTGCGAUCTAAACCTGTTCCGTGTTTGTGG
R-6832	TCTGCGAUCTAAGCGGGGTTCGGCCC
R-6849	TCTGCGAUTCATGCCGTGCGGTGGG
R-6861	TCTGCGAUTTATTTTCGCATCCTCCCCAAAC
R-6884	TCTGCGAUTCATGGAGGCAGATTCCAAATG
R-6889	TCTGCGAUCTAAGGCTGTATAAACTGCGATGGCAGG
R-6900	TCTGCGAUTTACAGGCGGGCATTGAC
R-6960	TCTGCGAUCTATCCGTGCTTTTCGTATCC
R-10763	TCTGCGAUTTATCCCTTGTTAATCCAACCG
R-10789	TCTGCGAUCTATTACCTAACAAGGGCAAAC
R-10809	TCTGCGAUCAGGTAAAAAGCCAAATTGATG

Continued on Next Page

Table 6.4 – Continued

Primer name	Sequence
R-11918	TCTGCGAUTCATCTAAAGGCTGTTGGCATAG
R-12045	TCTGCGAUCTAGGACGTCGAATCGCATG

Table 6.5: Plasmids used and constructed in this study. All plasmids are *ori^R** *amp^R*. *hph** contains a silent point mutation eliminating an AsiSI restriction site. *ori^R** contains a point mutation eliminating a Nb.BtsI nicking site. UEC = AsiSI/Nb.BtsI uracil excision cassette

Plasmid	Genotype	Source
AN0103	TS1::P <i>O</i> 7::AN0103::GFP-Stag::TtrpC::argB::TS2	This study
AN0129	TS1::P <i>O</i> 7::AN0129::GFP-Stag::TtrpC::argB::TS2	This study
AN0164	TS1::P <i>O</i> 7::AN0164::GFP-Stag::TtrpC::argB::TS2	This study
AN0410 (bimG)	TS1::P <i>O</i> 7::AN0410::GFP-Stag::TtrpC::argB::TS2	This study
AN0504 (sitA)	TS1::P <i>O</i> 7::AN0504::GFP-Stag::TtrpC::argB::TS2	This study
AN0914	TS1::P <i>O</i> 7::AN0914::GFP-Stag::TtrpC::argB::TS2	This study
AN1343	TS1::P <i>O</i> 7::AN1343::GFP-Stag::TtrpC::argB::TS2	This study
AN1358	TS1::P <i>O</i> 7::AN1358::GFP-Stag::TtrpC::argB::TS2	This study
AN1467	TS1::P <i>O</i> 7::AN1467::GFP-Stag::TtrpC::argB::TS2	This study
AN2472	TS1::P <i>O</i> 7::AN2472::GFP-Stag::TtrpC::argB::TS2	This study
AN2902	TS1::P <i>O</i> 7::AN2902::GFP-Stag::TtrpC::argB::TS2	This study
AN3793	TS1::P <i>O</i> 7::AN3793::GFP-Stag::TtrpC::argB::TS2	This study
AN3810	TS1::P <i>O</i> 7::AN3810::GFP-Stag::TtrpC::argB::TS2	This study
AN3941 (nimT)	TS1::P <i>O</i> 7::AN3941::GFP-Stag::TtrpC::argB::TS2	This study
AN4419	TS1::P <i>O</i> 7::AN4419::GFP-Stag::TtrpC::argB::TS2	This study
AN4426	TS1::P <i>O</i> 7::AN4426::GFP-Stag::TtrpC::argB::TS2	This study
AN4544	TS1::P <i>O</i> 7::AN4544::GFP-Stag::TtrpC::argB::TS2	This study
AN4896	TS1::P <i>O</i> 7::AN4896::GFP-Stag::TtrpC::argB::TS2	This study
AN5057	TS1::P <i>O</i> 7::AN5057::GFP-Stag::TtrpC::argB::TS2	This study
AN5722	TS1::P <i>O</i> 7::AN5722::GFP-Stag::TtrpC::argB::TS2	This study
AN6391 (pphA)	TS1::P <i>O</i> 7::AN6391::GFP-Stag::TtrpC::argB::TS2	This study
AN6892	TS1::P <i>O</i> 7::AN6892::GFP-Stag::TtrpC::argB::TS2	This study
AN6982	TS1::P <i>O</i> 7::AN6982::GFP-Stag::TtrpC::argB::TS2	This study
AN8820 (cnaA)	TS1::P <i>O</i> 7::AN8820::GFP-Stag::TtrpC::argB::TS2	This study

Continued on Next Page

Table 6.5 – Continued

Plasmid	Genotype	Source
AN10077	TS1::P <i>O</i> 7::AN10077::GFP-Stag::TtrpC::argB::TS2	This study
AN10138	TS1::P <i>O</i> 7::AN10138::GFP-Stag::TtrpC::argB::TS2	This study
AN10281	TS1::P <i>O</i> 7::AN10281::GFP-Stag::TtrpC::argB::TS2	This study
AN10570	TS1::P <i>O</i> 7::AN10570::GFP-Stag::TtrpC::argB::TS2	This study
RFP	TS1::P <i>O</i> 7::rfp::TtrpC::argB::TS2	This study
GFP	TS1::P <i>O</i> 7::gfp::TtrpC::argB::TS2	This study
AN5859	TS1::P <i>O</i> 7::AN5859::TtrpC::argB::TS2	This study
AN5894	TS1::P <i>O</i> 7::AN5894::TtrpC::argB::TS2	This study
AN5955	TS1::P <i>O</i> 7::AN5955::TtrpC::argB::TS2	This study
AN5966	TS1::P <i>O</i> 7::AN5966::TtrpC::argB::TS2	This study
AN6030	TS1::P <i>O</i> 7::AN6030::TtrpC::argB::TS2	This study
AN6062	TS1::P <i>O</i> 7::AN6062::TtrpC::argB::TS2	This study
AN6091	TS1::P <i>O</i> 7::AN6091::TtrpC::argB::TS2	This study
AN6166	TS1::P <i>O</i> 7::AN6166::TtrpC::argB::TS2	This study
AN6178	TS1::P <i>O</i> 7::AN6178::TtrpC::argB::TS2	This study
AN6195	TS1::P <i>O</i> 7::AN6195::TtrpC::argB::TS2	This study
AN6221	TS1::P <i>O</i> 7::AN6221::TtrpC::argB::TS2	This study
AN6223	TS1::P <i>O</i> 7::AN6223::TtrpC::argB::TS2	This study
AN6293	TS1::P <i>O</i> 7::AN6293::TtrpC::argB::TS2	This study
AN6295	TS1::P <i>O</i> 7::AN6295::TtrpC::argB::TS2	This study
AN6325	TS1::P <i>O</i> 7::AN6325::TtrpC::argB::TS2	This study
AN6385	TS1::P <i>O</i> 7::AN6385::TtrpC::argB::TS2	This study
AN6420	TS1::P <i>O</i> 7::AN6420::TtrpC::argB::TS2	This study
AN6492	TS1::P <i>O</i> 7::AN6492::TtrpC::argB::TS2	This study
AN6508	TS1::P <i>O</i> 7::AN6508::TtrpC::argB::TS2	This study
AN6536	TS1::P <i>O</i> 7::AN6536::TtrpC::argB::TS2	This study
AN6549	TS1::P <i>O</i> 7::AN6549::TtrpC::argB::TS2	This study
AN6569	TS1::P <i>O</i> 7::AN6569::TtrpC::argB::TS2	This study
AN6575	TS1::P <i>O</i> 7::AN6575::TtrpC::argB::TS2	This study
AN6650	TS1::P <i>O</i> 7::AN6650::TtrpC::argB::TS2	This study
AN6687	TS1::P <i>O</i> 7::AN6687::TtrpC::argB::TS2	This study
AN6696	TS1::P <i>O</i> 7::AN6696::TtrpC::argB::TS2	This study
AN6715	TS1::P <i>O</i> 7::AN6715::TtrpC::argB::TS2	This study

Continued on Next Page

Table 6.5 – Continued

Plasmid	Genotype	Source
AN6733	TS1::P <i>O</i> 7::AN6733::TtrpC::argB::TS2	This study
AN6762	TS1::P <i>O</i> 7::AN6762::TtrpC::argB::TS2	This study
AN6788	TS1::P <i>O</i> 7::AN6788::TtrpC::argB::TS2	This study
AN6814	TS1::P <i>O</i> 7::AN6814::TtrpC::argB::TS2	This study
AN6828	TS1::P <i>O</i> 7::AN6828::TtrpC::argB::TS2	This study
AN6849	TS1::P <i>O</i> 7::AN6849::TtrpC::argB::TS2	This study
AN6900	TS1::P <i>O</i> 7::AN6900::TtrpC::argB::TS2	This study
AN10763	TS1::P <i>O</i> 7::AN10763::TtrpC::argB::TS2	This study
AN10789	TS1::P <i>O</i> 7::AN10789::TtrpC::argB::TS2	This study
AN10809	TS1::P <i>O</i> 7::AN10809::TtrpC::argB::TS2	This study
AN11918	TS1::P <i>O</i> 7::AN11918::TtrpC::argB::TS2	This study
AN12045	TS1::P <i>O</i> 7::AN12045::TtrpC::argB::TS2	This study

6.5 Chemical analysis of strains

Analysis was performed using UPHLC-DAD-TOFMS. The equipment used was an Agilent 6550 iFunnel Q-TOF LC/MS system (Agilent Technologies, Torrence, CA, USA). This system comprised an Agilent 6550 orthogonal acceleration quadrupole time of flight mass spectrometer (TOFMS) equipped with an iFunnel ESI source, connected to an Agilent 1290 infinity UHPLC. The column used was an Agilent Poroshell 120 phenyl hexyl 2.7 μ m, 250 mm x 2.1 mm, and the column was maintained at 60 °C. A linear water-acetonitrile (LCMS-grade) gradient was used (both solvents were buffered with 20 mM formic acid) starting from 10 % (vol/vol) acetonitrile and increased to 100 % in 15 min, maintaining this ratio for 2.5 min before returning to starting conditions in 0.1 min and staying there for 2.4 min before the next run. A flow rate of 0.35 ml/min was used. TOFMS was performed in ESI+ with a data acquisition rate of 10 scans per second at m/z 30-1700. The TOFMS was calibrated before each run sequence using a proprietary Agilent calibration, using the Agilent ESI-L tuning mix. During analysis the lockmass solution, containing compounds forming ions with m/z 186.2216 and m/z 922.0100, was constantly infused into the spray chamber using the iFunnel. UV/VIS spectra were collected at wavelengths from 190 to 640 nm. Data processing was performed using Agilent MassHunter B06.00. Tandem MS was performed with a fragmentation energy of 10, 20, and 40 eV, respectively.

Chapter 7

PHUSER (Primer Help for USER)

This chapter contains the published paper "PHUSER (Primer Help for USER): a novel tool for USER fusion primer design", published in *Nucleic Acids Research*, of which I am a co-author.

In this study we have developed a web-based program for designing primers that are compatible with the uracil-excision (UE) cloning method. This program helps design primers for constructs to be inserted into any UE-cassette, fusion of different sequences with subsequent insertion into any UE-cassette, or fusion of several sequences by UE-compatible overhangs but without inserting the sequence into a vector.

I have used this program to design primers for amplification of *A. niger* PKS genes (see Chapter 8) for heterologous expression in *A. nidulans*, some of which were used for testing the PHUSER program. Furthermore, I have used the program to design primers for amplification of *A. nidulans* phosphatases and transcription factors, used for tunable over-expression at IS1 in *A. nidulans* (see Chapter 6).

My contribution to this study included assessment of program contents and user-friendliness, testing usability of the PHUSER program, and experimental validation of primer fitness. Supporting information is included in Appendix C.

PHUSER (Primer Help for USER): a novel tool for USER fusion primer design

Lars Rønn Olsen^{1,*}, Niels Bjørn Hansen², Mads Tvillinggaard Bonde²,
Hans Jasper Genée², Dorte Koefoed Holm², Simon Carlsen², Bjarne Gram Hansen²,
Kiran Raosaheb Patil³, Uffe Hasbro Mortensen^{2,*} and Rasmus Wernersson^{1,*}

¹Center for Biological Sequence Analysis, Department of Systems Biology, Technical University of Denmark, Building 208, ²Center for Microbial Biotechnology, Department of Systems Biology, Technical University of Denmark, Building 223, DK-2800 Lyngby, Denmark and ³Structural and Computational Biology Unit, European Molecular Biology Laboratory, Meyerhofstrasse 1, 69117, Heidelberg, Germany

Received February 26, 2011; Revised April 19, 2011; Accepted May 3, 2011

ABSTRACT

Uracil-Specific Exision Reagent (USER) fusion is a recently developed technique that allows for assembly of multiple DNA fragments in a few simple steps. However, designing primers for USER fusion is both tedious and time consuming. Here, we present the Primer Help for USER (PHUSER) software, a novel tool for designing primers specifically for USER fusion and USER cloning applications. We also present proof-of-concept experimental validation of its functionality. PHUSER offers quick and easy design of PCR optimized primers ensuring directionally correct fusion of fragments into a plasmid containing a customizable USER cassette. Designing primers using PHUSER ensures that the primers have similar annealing temperature (T_m), which is essential for efficient PCR. PHUSER also avoids identical overhangs, thereby ensuring correct order of assembly of DNA fragments. All possible primers are individually analysed in terms of GC content, presence of GC clamp at 3'-end, the risk of primer dimer formation, the risk of intra-primer complementarity (secondary structures) and the presence of polyN stretches. Furthermore, PHUSER offers the option to insert linkers between DNA fragments, as well as highly flexible cassette options. PHUSER is publicly available at <http://www.cbs.dtu.dk/services/phuser/>.

INTRODUCTION

The assembly of several DNA fragments is a common task in molecular biology and traditionally this has been achieved by merging matching ends generated by restriction endonucleases. However, this rarely allows for a seamless fusion of two fragments, which is often desirable. Moreover, the need for unique restriction sites at the site of desired fusion can render the process practically unfeasible for large DNA segments; and if possible, the process is time consuming and tedious. A popular way to bypass these obstacles is to merge DNA fragments by fusion PCR (also known as overlapping PCR or overlap extension PCR) (1). While this technique is suitable for achieving seamless fusion, it requires PCR optimization and the amplification rate drops dramatically with an increase in the number of fragments to be fused and/or when the final product size is >4 kb (2,3). USER fusion, which allows rapid and efficient assembly of multiple PCR fragments in a one-step approach (4,5), offers a number of advantages over traditional cloning procedures and fusion PCR: (i) it is independent of restriction sites; (ii) it does not require DNA ligase; (iii) it eliminates the need for amplifying very large fragments and (iv) the final chimerical DNA fragment, which is inserted into the destination vector, is generated from sequences that have only been PCR amplified once, hence, reducing the amounts of PCR generated errors that often compromise fragments fused by PCR. The only requirement of USER fusion is the presence of an A and a T separated by 8–20 bp's in the region that one wishes to fuse (see Figure 1 for an overview of the process). This region is hereafter referred to as segment. The design of the primers for

*To whom correspondence should be addressed. Tel: +1 617 582 7951; Fax: +1 617 632 3351; Email: lronn@bio.dtu.dk
Correspondence may also be addressed to Uffe Mortensen. Tel: +45 45252701; Fax: +45 45884148; Email: um@bio.dtu.dk
Correspondence may also be addressed to Rasmus Wernersson. Tel: +45 45252489; Fax: +45 45931585; Email: raz@cbs.dtu.dk

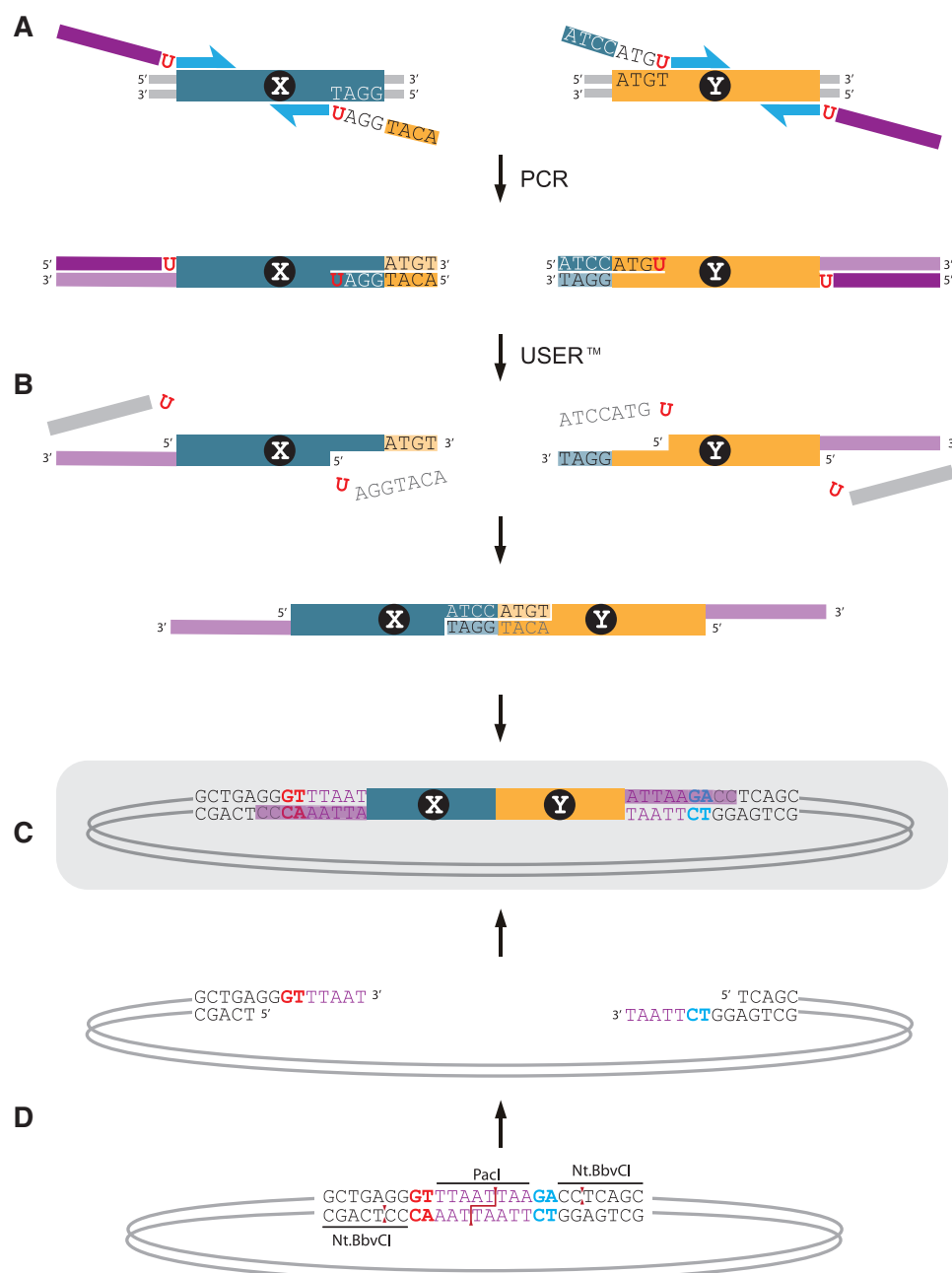


Figure 1. Overview of the USER fusion process. (A) PCR amplification of target DNA using uracil containing primers and a proofreading polymerase which stalls at uracil during PCR. (B) The PCR products are treated with the uracil DNA glycosylase and DNA glycosylase-lyase Endo VIII enzyme mix (USER™), creating 3' extensions. The primer design allows complementary hybridization of overhangs between fragments. (C) The overhangs (purple) match extensions of a prepared USER compatible vector. Complementary base pairing facilitates self-assembly of the fragments and the vector. The construct is now ready for transformation in *E. coli*. (D) Preparation of USER compatible vector is done by digestion of the USER cassette with restriction and nicking enzymes.

PCR amplification of the DNA fragments is based on these segments, and after PCR amplification the DNA fragments are treated with USER™ reagent and will self-assemble in a pretreated plasmid containing a USER cassette. The independence from restriction sites, along with the fact that USER fusion does not require subsequent rounds of PCR (as fusion PCR does), makes the method particularly favourable.

Given that USER fusion is based on a single-step cloning approach, the design of good primers is essential

for ensuring correct and efficient fusions. Designing primers for USER applications is a tedious and time consuming task to perform manually, due to the number of factors (e.g. defining the USER tails, making compatible ends that can be joined in an orderly fashion, ensuring uniform T_m , avoiding hairpins and primer dimer formation), which must be taken into account. To address this problem, we have developed the PHUSER (Primer Help for USER) primer design software that employs a simple, yet powerful algorithm to determine the best primers for

USER fusion, evaluated based on a number of parameters described in the 'Results page' section of this article. PHUSER is user friendly and significantly simplifies the primer design step of USER fusion of multiple DNA fragments.

SOFTWARE FEATURES

PHUSER is available as an easy-to-use web server aimed at molecular biologists, as well as a downloadable UNIX command-line tool aimed at more technically oriented users to whom a scriptable interface to the program is desirable. The web server has been tested in Internet Explorer, Firefox and Safari on Mac OSX, Linux and Windows. The command-line tool is written in Perl with a Python wrapper, and will thus be compatible with most existing computer platforms.

Primer design

The input data for the program are any number of DNA fragments in FASTA format, and a USER cassette consisting of a restriction site and a nicking site, which matches the cassette to be used in the destination vector. Based on this input, PHUSER will identify all segments in proximity of the joining region [i.e. the junction between two DNA fragments (4)] of the DNA fragments containing a dTTP located 8–20 bases downstream of an dATP. This typically gives rise to a number of joining regions, which will produce overhang in the length of 8–20 bases, also called fusion tails. When fusing more than two DNA fragments, unique fusion tails must be used to ensure that the fusion of fragments occurs in the correct order. There is no computational limit as to how many fragments can be processed using PHUSER, although the number of unique fusion tails may be exhausted if multiple fragments are submitted. Six fragments have been reported successfully fused in a single process (6), but the practical experimental limitations have yet to be established.

The part of the primer that binds to the target DNA is here referred to as the binding region. The desired length of a primer binding region is 18–24 bp (7), leading to seven potential primer lengths for both the forward and reverse strand. The selection of the pair of primers is based on matching the annealing temperature, T_m ; the T_m is calculated for the binding region part of the primer only. For a given primer pair, the primers' T_m should be within 2°C of each other, in order to achieve efficient PCR amplification (8). The T_m for every possible forward and reverse primer is calculated using the nearest neighbor method (9), and pairs that obey the 2°C rule are stored. If no primers obey the rule, the pairs that have the lowest T_m difference are processed further. In the event that fusion tails are identical, PHUSER compares all the primer pairs, disqualifies pairs with identical fusion tails and selects the next best T_m match as the best primer pair.

Cassette options

PHUSER features a choice of five predefined USER cassettes: (i) *PacI/Nt.BbvCI* (10); (ii) *XbaI(2)/Nt.*

BbvCI (<http://www.neb.com/nebecomm/ManualFiles/manualE5500.pdf>) (11); (iii) *AsiSI/Nb.BsmI* (12); (iv) *PmeI/Nb.BbvCI*; and (v) *AsiSI/Nb.BtsI* (12). Linear assembly (12) is also supported, when PHUSER is used without selecting a USER cassette. Furthermore, users have the option to tailor their own cassettes (see 'Query page' section).

Linker insertion

Lastly, PHUSER enables the insertion of linkers between the DNA fragments. This is done simply by including a FASTA entry between the fragments with the header '>linker'.

WEB INTERFACE FOR PHUSER 1.0

Query page

The PHUSER query page is designed to be simple and intuitive with a tabbed-based interface separating basic from advanced options. DNA sequences can be submitted in FASTA format either by copy/pasting or uploading a text file containing the sequences. PHUSER has a built-in feature to divide large fragments into several shorter fragments for PCR. The desired maximum length of the PCR products can be adjusted in the drop-down menu titled 'Automatic adjustment of maximum PCR product length'. If this option is used, PHUSER will automatically split the input sequence into a number of subsequences each below the defined threshold and design fusion primers for them all. This makes it possible to input a single large sequence (which would be considered too long for efficient PCR) and let the program take care of the entire process of designing primers for USER fusion for this particular sequence.

PHUSER offers three options for cassette selection, one basic and two advanced. The tab titled (i) 'Basic: Predefined fusion cassettes' allows for selection of one of the five predefined cassettes described above. The tab titled (ii) 'Advanced: Select restriction/nicking pair' allows users to custom design a cassette by selecting restriction and nicking enzymes, and entering directional bases. The tab titled (iii) 'Advanced: User designed cassettes' allows users to enter a custom cassette with the format:

5'-GCTGAGGGTTTAAT.TAAGACC.TCAGC-3'

The first period indicates the position of the restriction site and the second period indicates the position of the nicking site. This needs to be entered for both the forward and the reverse strand of the custom cassette.

Results page

Once the desired inputs are submitted, the user will be presented with the results page (Figure 2). The results report is divided into four sections:

- (i) Overview of the primers (5'–3'),
- (ii) Graphic overview of the DNA fragments, primers and their interaction,

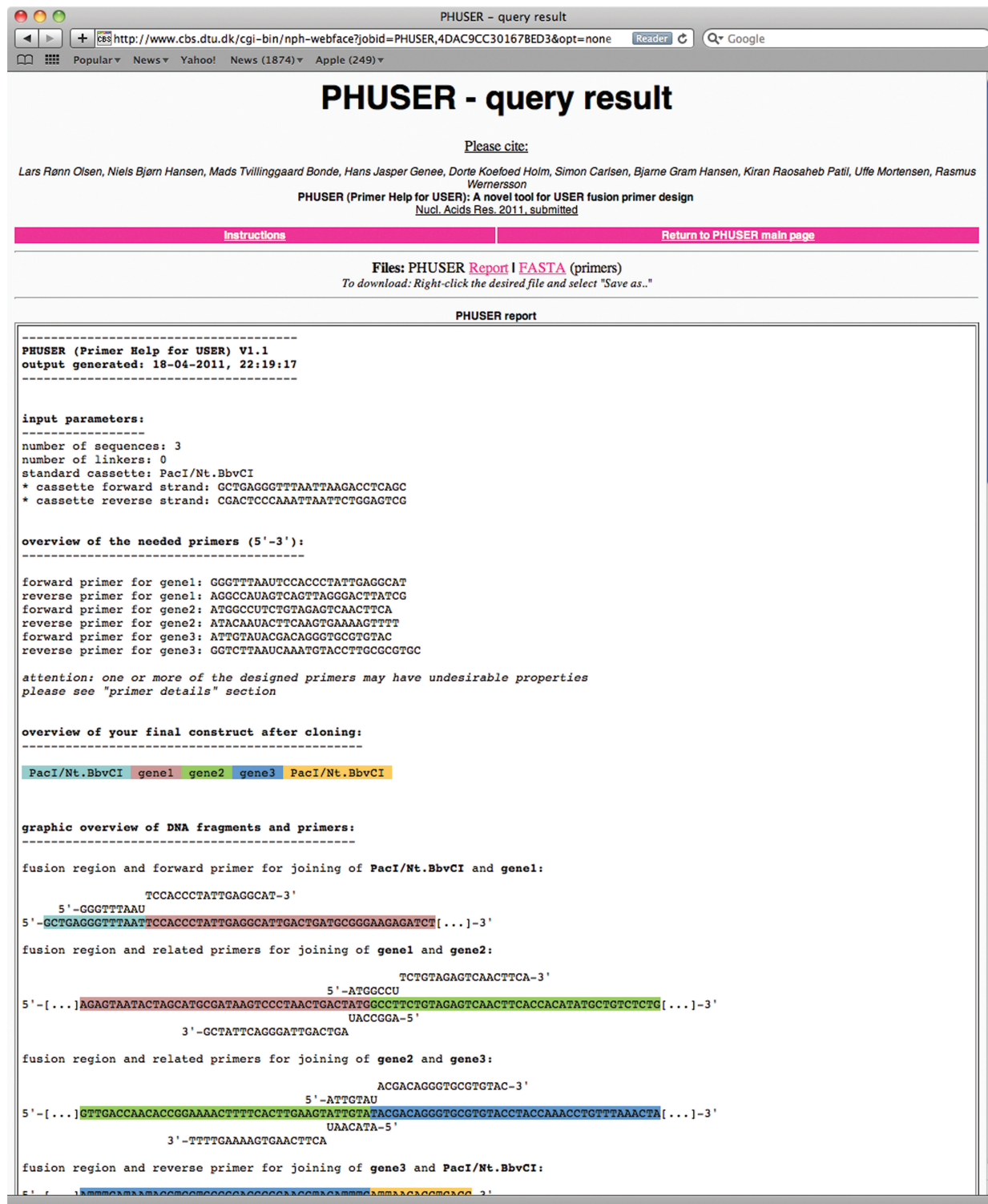


Figure 2. Result page of the PHUSER web-server. Here a detailed colour-coded report summarizing the main finding of the primer design is shown, and the user has the option of downloading the primers in FASTA format as well.

- (iii) Primer details including assessment of the parameters presented in the 'Primer design' section and
- (iv) The final construct after USER fusion.

The quick overview of the primers is written in 5' to 3' direction and includes the dUTP. The graphical overview

of DNA fragments and primers provides a visual presentation of where the primers anneal to the template sequences. This is followed by the primer details, showing more specific information regarding each primer. Apart from T_m , this section includes assessment of the following parameters.

primer details:

primer details for genel

```

forward primer: GGGTTTAAUTCCACCCTATTGAGGCAT
* Tm: 56.3C - in optimal region (55-72)? ...YES
* GC ratio is: 50.00% - in optimal region (40-60)? ...YES
* GC clamp present at 3' end? ...YES
* more than 3 G/C out of last 5 bases at 3' end? ...YES
* risk of primer dimer formation in primer pair? ...NO
* risk of intra-primer complementarity? ...NO
* presence of polyN stretches? ...NO

reverse primer: AGGCCAUAGTCAGTTAGGGACTTATCG
* Tm: 54.6C - in optimal region (55-72)? ...NO
* GC ratio is: 45.00% - in optimal region (40-60)? ...YES
* GC clamp present at 3' end? ...YES
* more than 3 G/C out of last 5 bases at 3' end? ...NO
* risk of primer dimer formation in primer pair? ...NO
* risk of intra-primer complementarity? ...NO
* presence of polyN stretches? ...NO

Tm of primers within 2C of each other? ...YES

```

Figure 3. Primer details from the PHUSER report. The desired properties are colour-coded in green and undesired properties are encoded in red.

GC content. The GC content (number of dGTP plus dCTP as a percentage of the total number of nucleic acid residues) should ideally be between 40% and 60%, to keep annealing temperature within a favourable range (9).

Presence of 3'-end GC clamp. Due to the stronger annealing of guanine and cytosine with complementary bases, specific binding is promoted by the presence of at least one of either of these bases in the last five bases of the 3'-end (GC-clamp). However, if more than three out of the last five bases at the 3'-end are guanine or cytosine, the binding to the DNA fragment may be too strong (8).

Risk of Inter- and intra-primer complementarity. Stretches of more than five complementary bases between different primers used in the same PCR can lead to primer dimer formation, which can interfere with the desired hybridization. Similarly, stretches of more than three complementary bases within a primer can lead to formation of intra-primer secondary structures, resulting in reduced PCR efficiency. With PHUSER the risk of inter- and intra-primer complementarity is assessed based on the relationship between the T_m of the primer and the complementary bases potentially leading to undesirable dimerization or intra-primer secondary structures. If the T_m of the potential inter-primer dimerization or intra-primer secondary structure formation is at least 10°C smaller than that of the annealing temperature used in the PCR cycle, the risk of complementarity is considered theoretical, since most of the DNA which could undesirably hybridize is denatured at this higher temperature.

Presence of polyN stretches. Sequences consisting of four or more consecutive identical nucleotides within the

primer are known as polyN stretches; polyG or polyC stretches promote non-specific annealing (8), whereas polyA or polyT can cause the 'opening' of stretches of the primer template complex, referred to as 'breathing' (8).

The assessment of each parameter is highlighted in either green (signifying a positive assessment) or red (signifying a negative assessment)—see Figure 3 for an example. Note that the primer parameters used for primer design are supplied only to evaluate the efficiency of a specific amplification and not the competence of the amplification. Therefore, the lack of positive assessment in any parameter (or all parameters) does not preclude a useful PCR outcome.

The last section shows the final product, which will be the result of assembling the sequences into the cassette using USER fusion. The user also has the option to download text files of the report as shown on the screen, or only the best primers in FASTA format, which can be directly used to order the primers.

PROOF OF CONCEPT

Polyketide synthases (PKSs) are complex enzymes encoded by large genes (~10 kb), and PCR amplification of genes of such lengths is very challenging. One way to overcome this problem is to amplify sections of PCR friendly sizes, followed by fragment reassembly to attain the original gene. USER fusion is a powerful method for assembly of such large genes. In this study, USER fusion was applied for the cloning of six genes encoding putative PKSs of known sequence from *Aspergillus niger* (Table 1). As proof of concept of PHUSER, the software was used for the design of primers (Supplementary Table S1).

According to size, the six PKS genes were amplified as 2, 3, 4 or 5 fragments generating a total of 21 PCR

Table 1. Overview of strategy and efficiency of cloning six PKSs

Model ID (JGI)	Gene annotation	No. of fragments cloned (excluding vector backbone)	Gene length (bp)	Length of PCR fragments cloned (bp)	Fraction correct clones ^a
44965	Putative PKS	2	5340	2624, 2740	2/4
56896	Putative PKS	3	6755	2250, 2170, 2260	2/4
128601	Putative PKS	5	11,657	2040, 2240, 2310, 2730, 2537	2/4
194381	Putative PKS	3	6955	2259, 2450, 2255	4/4
211885	Putative PKS	4	7996	1776, 2310, 1820, 2112	3/4
225574	<i>fumI</i>	4	8100	1620, 2310, 2030, 2027	2/4

The genes were amplified in fragments of 1.7–2.7kb and all primers were designed with PHUSER.

^aResult obtained by restriction analysis.

products. Prior to transformation into chemically competent *E. coli*, the fragments of the individual genes (five fragments at the most) were purified and mixed in equal quantities with the prepared cloning vector pU1111-1 and treated with USERTM mix. pU1111-1 is a 12kb expression vector for *Aspergillus nidulans*, including a *gpdA* promoter, an *AsiSI*/*Nb.BtsI* USER cassette, *TrpC* terminator and an *argB* marker flanked by UP- and DOWN-stream sequences for targeted insertion (12). Transformants were selected on LB media containing 100 µg/ml ampicillin, and the plasmids were purified and tested by restriction analysis. Despite the complexity of the six vector construction reactions, we found that in all cases, 50% or more of the transformants tested contained a plasmid with the correct configuration of inserts (Table 1). Furthermore, a selection of these plasmids was sequenced (StarSeq, Germany) confirming error-free sequence and assembly of all six genes. The false positives observed contained no insert and were therefore caused by incompletely digested pU1111-1. In other USER cloning experiments, we have shown that such false positives can be avoided by prolonged digestion with *AsiSI*/*Nb.BtsI*. For further details, please see Supplementary Data. The experiment demonstrates that the use of PHUSER for primer design results in a highly efficient cloning procedure.

Experimental details of the proof of concept study are found in the Supplementary Data.

CONCLUSION

USER fusion is considered an attractive alternative to classical restriction–ligation-based cloning, since it offers simplicity, speed and high efficiency. However, the laborious nature of manually designing good primers is a bottleneck in creating fusion constructs (10,13–15). PHUSER makes the design of good primers quick and easy, and will reduce the time needed to get from strategy to result significantly. Consequently, PHUSER in combination with USER fusion offer means to advance high-throughput generation of constructs. This allows for, e.g. simple construction of gene expression libraries and for vectors expressing genes encoding proteins that are fused to an epitope-tag, a purification-tag or to a fluorescent protein-like GFP. We have shown that PHUSER can be used to design primers that have been

experimentally shown to be efficient, by successful assembly of up to five PCR fragments in a single reaction with a success rate of 50% or higher for the six genes tested. PHUSER also supports different usages of the USER based cloning method, such as homologous recombination for library construction (13) and linear template construction (14). All these features render PHUSER not only an efficient, but also a very flexible tool for designing primers for USER fusion.

SUPPLEMENTARY DATA

Supplementary data are available at NAR Online.

ACKNOWLEDGEMENTS

We are thankful to Bo Salomonsen, Rasmus John Normand Frandsen and Annette Sørensen for beta-testing and constructive feedback.

FUNDING

Funding for scientific work: i) Computer resources (software development and hosting of the PHUSER web sever) was funded by a grant from the Danish Center for Scientific Computing. ii) Experimental work was funded by grants from the Danish Agency for Science, Technology and Innovation (grants 09-064967 and 09-064240 to K.R.P. and U.H.M.). Funding for open access charge: A grant from the Danish Center for Scientific Computing.

Conflict of interest statement. None declared.

REFERENCES

- Yon,J. and Fried,M. (1989) Precise gene fusion by PCR. *Nucleic Acids Res.*, **17**, 4895.
- Pont-Kingdon,G. (1997) Creation of chimeric junctions, deletions, and insertions by PCR. *Methods Mol. Biol.*, **67**, 167–172.
- Kuwayama,H., Obara,S., Morio,T., Katoh,M., Urushihara,H. and Tanaka,Y. (2002) PCR-mediated generation of a gene disruption construct without the use of DNA ligase and plasmid vectors. *Nucleic Acids Res.*, **30**, E2.
- Geu-Flores,F., Nour-Eldin,H.H., Nielsen,M.T. and Halkier,B.A. (2007) USER fusion: a rapid and efficient method for simultaneous fusion and cloning of multiple PCR products. *Nucleic Acids Res.*, **35**, e55.

5. Nour-Eldin,H.H., Geu-Flores,F. and Halkier,B.A. (2010) USER cloning and USER fusion: the ideal cloning techniques for small and big laboratories. *Methods Mol. Biol.*, **643**, 185–200.
6. Bitinaite,J., Rubino,M., Varma,K.H., Schildkraut,I., Vaisvila,R. and Vaiskunaite,R. (2007) USER friendly DNA engineering and cloning method by uracil excision. *Nucleic Acids Res.*, **35**, 1992–2002.
7. Dieffenbach,C.W., Lowe,T.M. and Dveksler,G.S. (1993) General concepts for PCR primer design. *PCR Methods Appl.*, **3**, S30–37.
8. Apte,A. and Saurabha,D. (2003) PCR primer design. In Dieffenbach,C.W. and Dveksler,G.S. (eds), *PCR Primer a Laboratory Manual*, 2 edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, USA.
9. SantaLucia,J. Jr (1998) A unified view of polymer, dumbbell, and oligonucleotide DNA nearest-neighbor thermodynamics. *Proc. Natl Acad. Sci. USA*, **95**, 1460–1465.
10. Nour-Eldin,H.H., Hansen,B.G., Norholm,M.H., Jensen,J.K. and Halkier,B.A. (2006) Advancing uracil-excision based cloning towards an ideal technique for cloning PCR fragments. *Nucleic Acids Res.*, **34**, e122.
11. Biolabs,N.E. (2008) USER™ Friendly Cloning Kit, (A Novel Tool for Cloning PCR Products by Uracil Excision) Catalog #E5500S, Version 1.3, 1/08.
12. Hansen,B.G., Salomonsen,B., Nielsen,M.T., Nielsen,J.B., Hansen,N.B., Nielsen,K.F., Regueira,T.B., Nielsen,J., Patil,K.R. and Mortensen,U.H. (2011) Versatile enzyme expression and characterization system for *Aspergillus nidulans*, with the Penicillium brevicompactum polyketide synthase gene from the mycophenolic acid gene cluster as a test case. *Appl. Environ. Microbiol.*, **77**(9), 3044–3051.
13. Villiers,B.R., Stein,V. and Hollfelder,F. (2010) USER friendly DNA recombination (USERec): a simple and flexible near homology-independent method for gene library construction. *Protein Eng. Des. Sel.*, **23**, 1–8.
14. Stein,V. and Hollfelder,F. (2009) An efficient method to assemble linear DNA templates for in vitro screening and selection systems. *Nucleic Acids Res.*, **37**, e122.
15. Geu-Flores,F., Olsen,C.E. and Halkier,B.A. (2009) Towards engineering glucosinolates into non-cruciferous plants. *Planta*, **229**, 261–270.

Chapter 8

Heterologous expression of polyketide synthases in *A. nidulans*

This study was the first study I conducted during my Ph.D., and it formed the basis for some of the other studies presented in this thesis. This study provided insights into *A. niger* secondary metabolism, which I have used for subsequent in-depth investigation of selected biosyntheses. Additionally, the study was used to assess the use of *A. nidulans* as a host for heterologous gene expression and reveal the obstacles in semi-HTP or HTP strain construction.

8.1 Introduction

The genome sequence and annotation of *A. niger* ATCC1015 was released to the public in 2006 [Baker, 2006], and is available from the DOE Joint Genome Institute database¹. Investigation of the genome had shown that the *A. niger* genome encoded no less than 37 putative PKS and PKS-like enzymes [Fisch et al., 2009], which was a number that greatly exceeded the number of known polyketide compounds from this species [Nielsen et al., 2009]. A very important observation was that the *A. niger* genome contained a gene cluster that was very similar to the fumonisin gene cluster in *Gibberella moniliformis* [Pel et al., 2007, Andersen et al., 2011, Proctor et al., 2003]. Further investigation of the metabolite profiles of different *A. niger* strains indeed verified that some strains of *A. niger* do produce the carcinogenic fumonisin compounds [Frisvad et al., 2007, 2013].

Fumonisin is produced by some industrial strains of *A. niger* [Frisvad et al., 2013], but the fact that *A. niger* has been used in industrial fermentations for many years without anyone reporting the presence of fumonisin, raises the question of whether or not this important workhorse is capable of producing any other harmful compounds that have not yet been reported. On the other hand, *A. niger* might be able to produce

¹<http://genome.jgi-psf.org/Aspni5/Aspni5.home.html>

new compounds that could be exploited for commercial purposes e.g. pigments or pharmaceuticals.

As many gene clusters are silenced under certain circumstances, it is very likely that some of these genes, if intact, are not expressed under standard laboratory or industrial cultivations. Based on this hypothesis, it should be possible to identify new compounds by heterologously expressing the putative PKS encoding genes under control of a constitutive promoter. In this study, I heterologously expressed the 37 putative PKS, PKS-like, and hybrid PKS/NRPS genes from *A. niger*. The PKSs were expressed in the model fungus *A. nidulans* using the characterized expression system developed by Hansen and co-workers [Hansen et al., 2011]. Using this system, the genes were placed under control of the strong, constitutive promoter *PgpdA* [Punt et al., 1990, 1991], and expressed from a constitutively expressed integration site, IS1 [Hansen et al., 2011].

A. niger is very well-studied with regard to production of secondary metabolites. Strains of *A. niger* primarily produces compounds that belong to the groups: fumonisins (PKs), naphtho- γ -pyrones (PKs), bicoumarins (PKs), malformins (NRPs), and asperazines (alkaloids) (selected compounds are shown in Figure 8.1) (reviewed by Nielsen et al. [2009]). *A. niger* produces very few terpenoid compounds, however compounds of the yanuthone family are produced in high amounts in the strain used in this study (see details in Chapter 10). During the past three years there have been a lot of advances within the field of linking genes to their corresponding metabolite(s) in *A. niger*. The metabolites that have been linked to genes are shown in Table 8.2.3, and the applied approaches have been discussed in Chapter 2.

8.2 Results and discussion

8.2.1 Heterologous expression of 37 putative PKSs links three PKSs to a compound

PKSs were chosen based on the genomic mining study performed by Fisch and co-workers [Fisch et al., 2009], an overview is shown in table 8.2.3. The genes were PCR amplified from KB1001 in one or two pieces, depending on the size of the gene. Generally, genes longer than 8 kb were split in two. The genes were inserted into the CMBU1111 plasmid [Hansen et al., 2011] which contains the constitutively active promoter *PgpdA*, terminator *TtrpC* and the *argB* marker, flanked by upstream (TS1) and downstream (TS2) targeting sequences for insertion into IS1 (Figure 8.2).

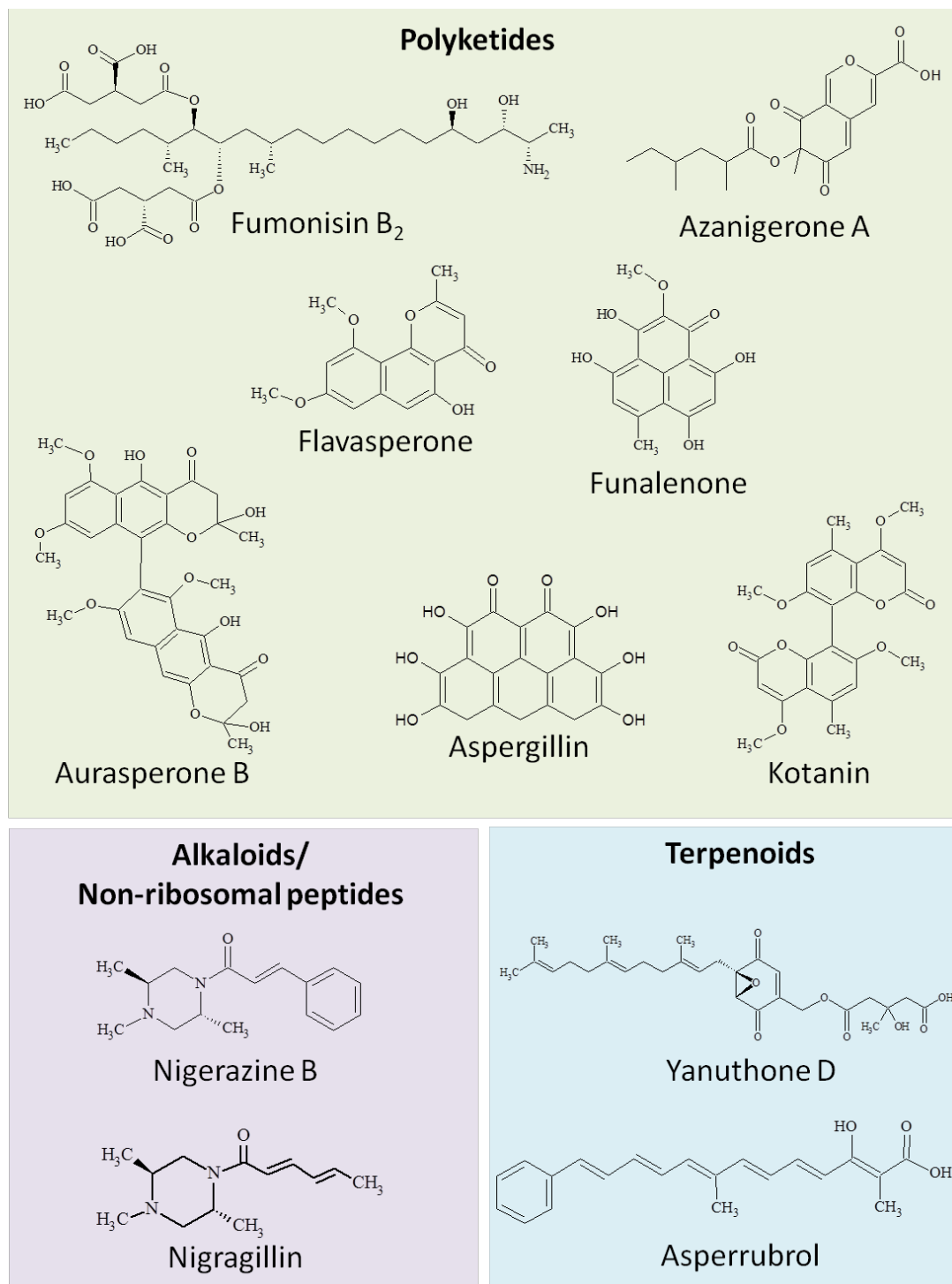


Figure 8.1: Selected *A. niger* secondary metabolites.

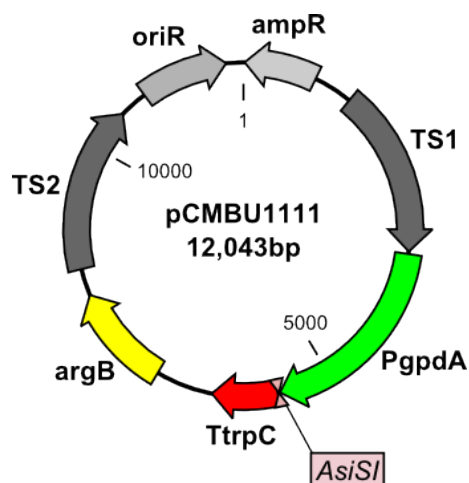


Figure 8.2: Cloning vector pCMBU1111.

All plasmids were verified by restriction analysis, linearized, and transformed into *A. nidulans* strain NID3 (contains a transient disruption of *nkuA*). Two of the resulting strains (ASPENIDRAFT_56896 (56896) and ASPENIDRAFT_191422 (191422)) had visible new phenotypes: 56896 produced vast amounts of a yellow compound and 191422 produced a slightly yellow pigment which was vaguely visible to the eye when looking at the fungus (see Figure 8.3). Moreover, it seemed that expression of both 191422 and 56896 resulted in darker conidia, which made the fungus appear larger compared to the reference where the outermost part of the colony is whitish transparent due to lack of conidial pigmentation.

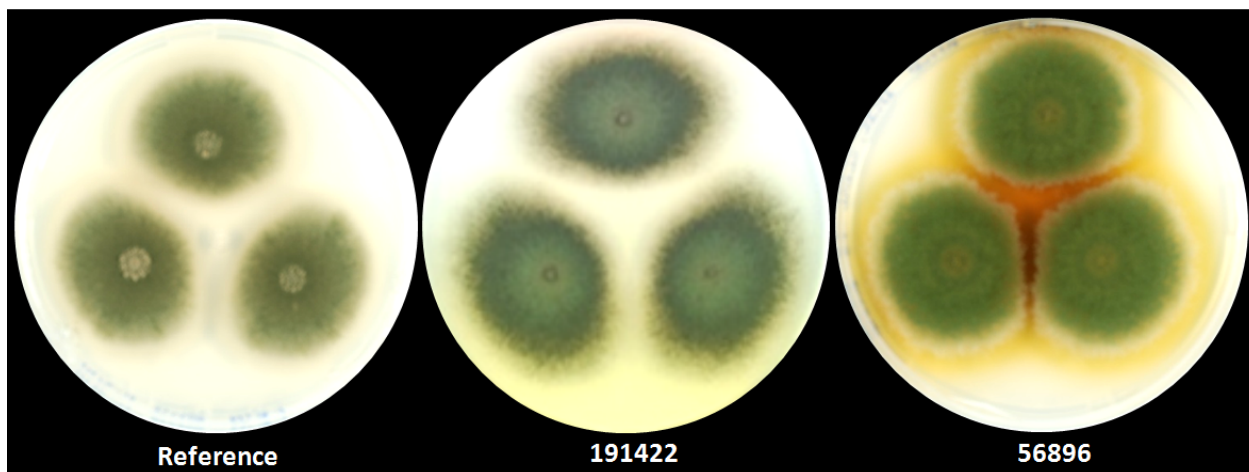


Figure 8.3: Phenotypes of *A. nidulans* expressing the ASPENIDRAFT_191422 and ASPENIDRAFT_56896 genes.

The metabolite profiles of all transformants were analyzed by HPLC-DAD-TOFMS. The results showed that two of the strains, expressing ASPENIDRAFT_44965 (44965) and 56896, produced a new compound compared to the reference (Figure 8.4). Interestingly, there was no detection of a new compound in the

strain expressing 191422, which displayed an altered phenotype (Figure 8.4).

Dereplication of the compound produced by 44965, and comparison to an authentic standard, confirmed that this PKS produced the small aromatic compound 6-methylsalicylic acid (6-MSA) (data shown in Chapter 10). The discovery of this compound was the beginning of uncovering the entire biosynthetic pathway towards the antibiotic and antifungal compound yanuthone D, which is described in chapter 10.

Dereplication of the compound produced by 56896 identified YWA1 as the PKS product. This compound was previously described for *A. nidulans* as the building block for green melanin, which is the green conidial pigment that gives *A. nidulans* its characteristic green color [Watanabe et al., 1999]. Examination of 56896 and 191422 later revealed that both of these PKSs produce YWA1, which is the precursor to naphtho- γ -pyrones and black melanin in *A. niger*. These findings are presented in detail in chapter 9.

8.2.2 Co-expression with discrete thioesterases did not provide any new compounds

An interesting observation for the PKSs that provided a product was that they all contained a release domain, and the remaining PKSs did not contain an obvious release domain (see Table 8.2.3). This domain is essential for release of the final compound from the PKS. Although it is not obvious, it was recently shown that the 6-MSA synthase contains a thiohydrolase (TH) domain rather than a dehydrogenase (DH) domain, and this TH domain is responsible for release of 6-MSA from the PKS [Moriguchi et al., 2010].

Analysis of the gene clusters revealed that three of the PKSs contained a discrete thioesterase domain within the proposed cluster: ASPNIDRAFT_44005 (44005), ASPNIDRAFT_179079, and ASPNIDRAFT_179585. The three thioesterases were co-expressed from a plasmid (pDHX2) in the strains harboring the respective PKSs. The strains were analyzed by HPLC-DAD-TOFMS but no new compounds were detected for any of the strains (data not shown).

8.2.3 Gene deletion links 44005 to asperrubrol

Puzzled, and intrigued, by the lack of new compounds I set out to investigate why no other compounds were produced. As a test case, I chose to delete 44005 in *A. niger* and examine the metabolite profile. Surprisingly, the 44005 Δ strain failed to produce asperrubrol - a compound which has so far been reported as a terpene (Figure 8.5) [Rabache et al., 1974, Nielsen et al., 2009].

Besides the thioesterase, examination of the genes surrounding 44005 revealed several activities including an NRPS. The presense of a putative NRPS in the proposed cluster was surprising, as asperrubrol does not contain any nitrogen atoms. Therefore, I also deleted the NRPS gene, but I did not detect any changes in the metabolite profile for the NRPS Δ strain (Figure 8.5).

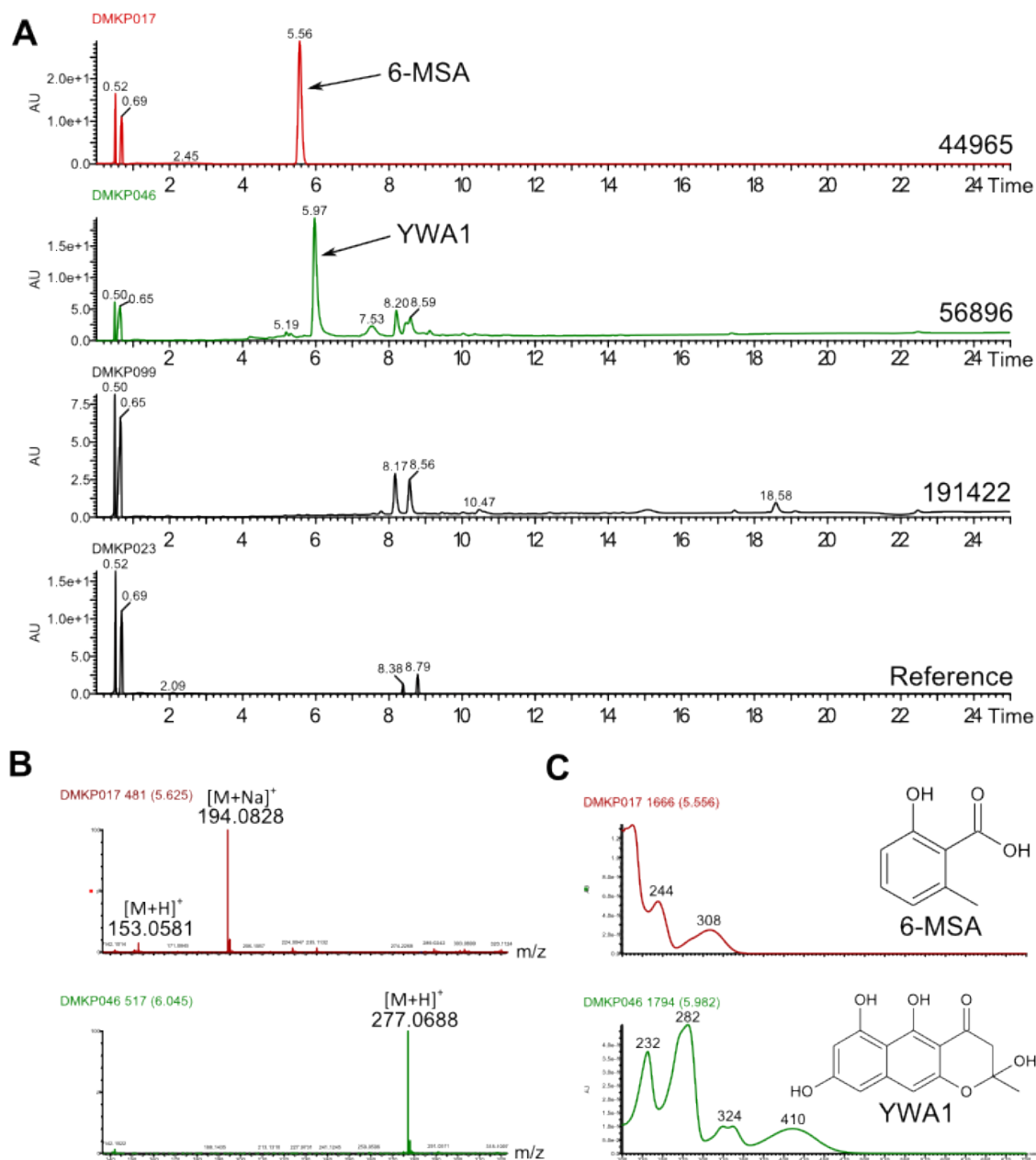


Figure 8.4: Dereplication of new compounds. **A.** DAD chromatogram of the strains expressing 44965, 56896, and 191422 in comparison to a reference. **B.** Mass spectra of the compound produced by 44965 (TOP) and of the compound produced by 56896 (BOTTOM). **C.** UV spectra of the compound produced by 44965 (TOP) and of the compound produced by 56896 (BOTTOM).

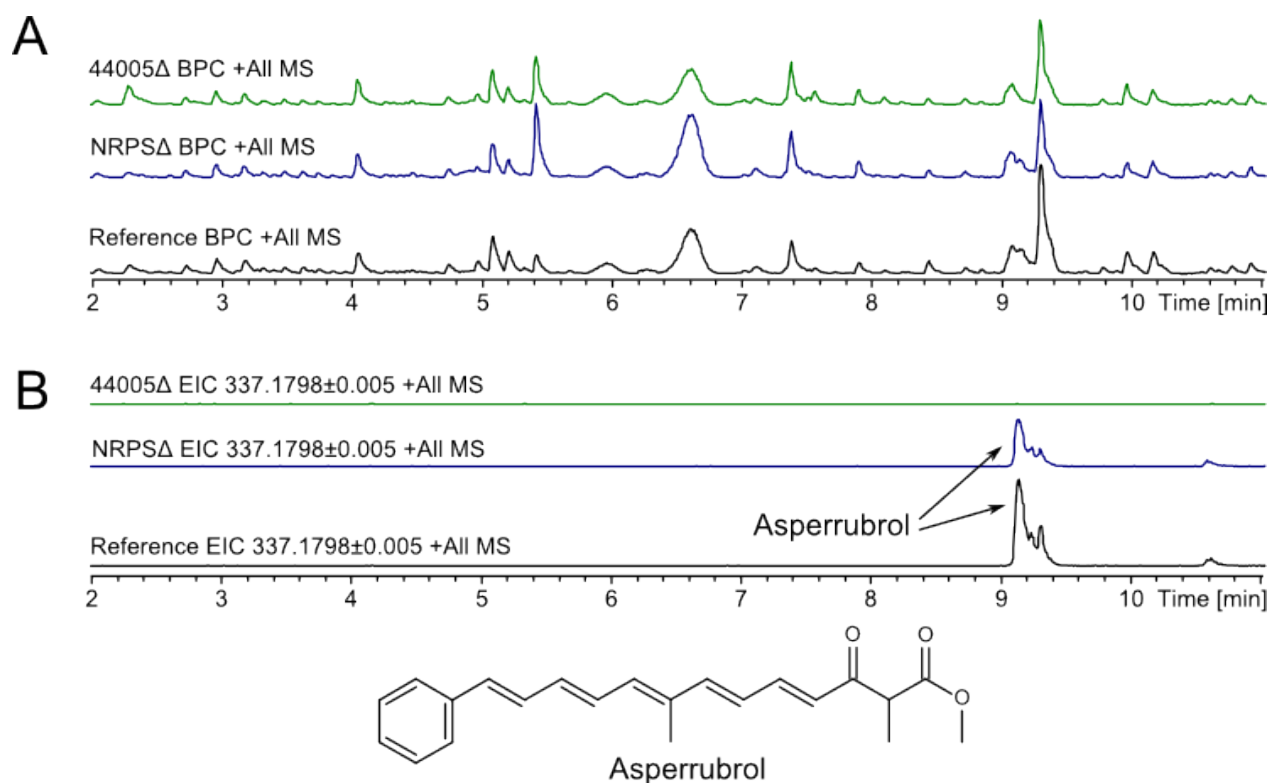


Figure 8.5: BPC and EIC chromatograms of the 44005Δ and the NRPSΔ strains. $[M+H]^+$ (asperrubrol) = 337.1798

I investigated the remainder of proposed cluster genes, and via BLAST analysis I found that 190482 had homology to amino acid permeases, which are enzymes involved in transport of amino acids into the cell. Furthermore, the BLAST analysis showed that 50432 is homologous to phenylalanine ammonia lyases (PALs), which convert phenylalanine into cinnamic acid. 44006 is homologous to 4-coumarate-CoA ligases, catalyzing attachment of Coenzyme A (CoA) to 4-coumarate, which is a compound that is identical to cinnamic acid, but with a hydroxyl on the aromatic ring.

Based on the enzyme activities encoded by the proposed gene cluster, I have suggested a model for biosynthesis of asperrubrol in *A. niger* (Figure 8.6). In this model I propose that the role of the putative amino acid permease (190482) is to transport phenylalanine across a plasmamembrane, be it into the cell or into a specific compartment of the cell. Next, phenylalanine is converted into cinnamic acid by the PAL (50432) and the cinnamic acid is attached to a CoA, catalyzed by a cinnamic acid ligase (44006). Cinnamic acid-CoA then serves as a starter unit to the PKS (44005), which performs five rounds of elongation while reducing three ketides to double bonds and attaching two methyl groups. The demethylasperrubrol compound is then released from the PKS by the thioesterase (44004). As there are no reports of O-methyl transferase activity within a PKS the final methyl group must be attached post synthesis, however this activity is not obviously encoded by any gene in the proposed cluster. The asperrubrol compound was identified by dereplication, and has not been compared to an authentic standard or investigated by NMR. Thus, it is

possible and even likely that the final methyl group is not an O-methyl group, but a C-methyl group that has been attached somewhere else on the PK chain by the PKS. There are many different reports on the structure of asperrubrol and the similar structure asperenone/aspergellone when comparing data from Antibase2010, Nielsen et al. [2009], and original literature e.g. [Rabache et al., 1974]. Thus I suspect that there are many different analogues of these compounds, and that they are strain specific within the group of black *Aspergilli*.

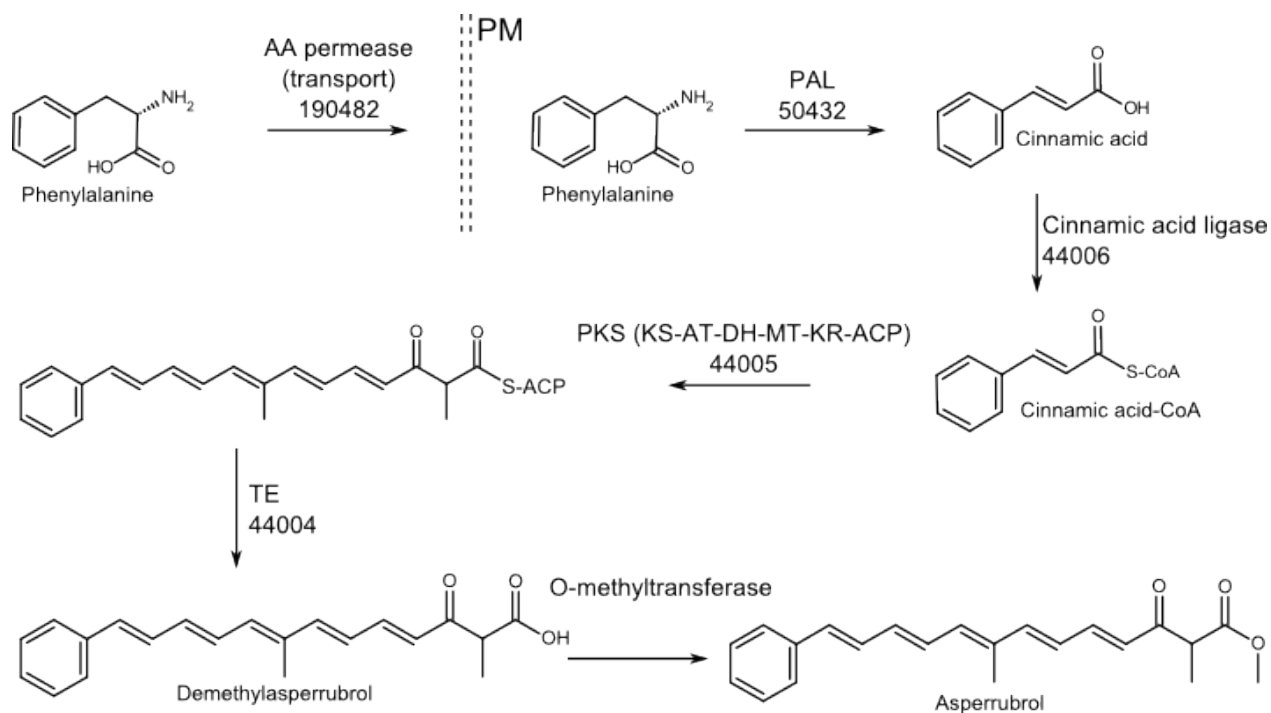


Figure 8.6: Proposed biosynthesis of asperrubrol in *A. niger*.

In order to test my model, I heterologously expressed the PAL (50432) in *A. nidulans* to verify that it did produce cinnamic acid. Inspection of the resulting metabolite profile and comparison to an authentic cinnamic acid standard confirmed that cinnamic acid was indeed the product of the PAL (Figure 8.7). As expected, the *A. nidulans* reference strain did not produce detectable amounts of cinnamic acid.

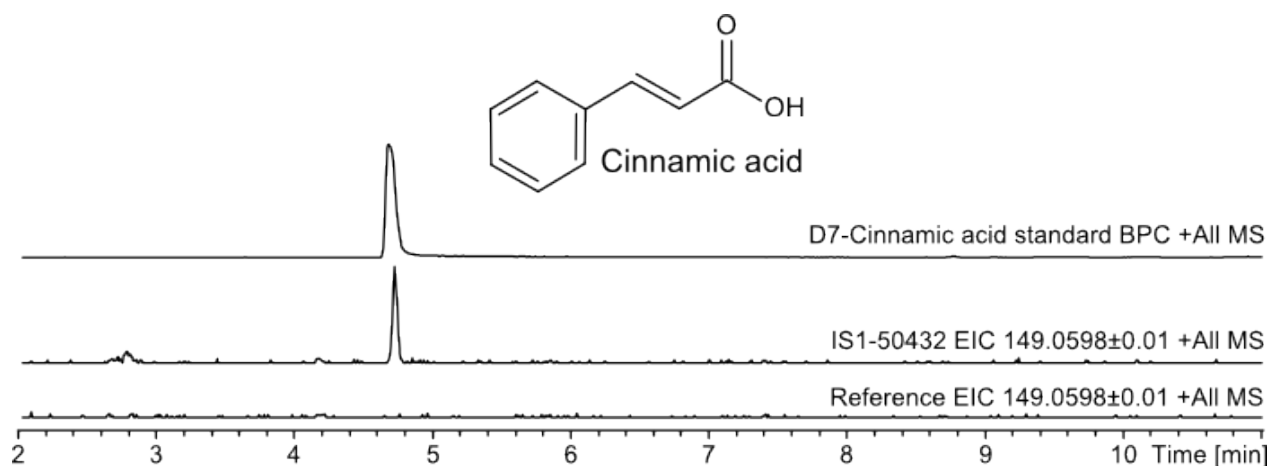


Figure 8.7: EIC chromatograms of IS1-50432 and reference compared to a cinnamic acid standard. $[M+H]^+$ (cinnamic acid) = 149.0598 Da.

Investigation of this gene cluster indeed clarified why heterologous expression of the PKS, and even in combination with the TE, did not provide any products. I have shown that the 44005 PKS is dependent on an alternative starter unit, cinnamic acid, which is not produced in *A. nidulans*. Additionally, I have linked a PKS to a known *A. niger* compound and importantly, I have found that a compound that was originally proposed to be of terpenoid origin, is in fact a PK that uses cinnamic acid (phenylalanine-derived) as a precursor.

This study is not representative of all the PKSs that did not produce a compound by heterologous expression, however, it does provide a valuable insight into the complexity of PK synthesis. It also indicates that some of the partly or highly reducing PKSs may in fact produce known compounds that have previously been regarded as terpenoids. Table 8.2.3 shows the putative and characterized PKSs in *A. niger*, and the compounds that have been linked to them as per August 2013 (including unpublished data obtained in this study).

Table 8.1: *A. niger* PKSs. *ASPNI DRAFT_#, ¹[Gil Girol et al., 2012]**, ²[Zabala et al., 2012], ³[Li et al., 2011b], ⁴[Pel et al., 2007, Frisvad et al., 2007, Khaldi and Wolfe, 2011]

Model ID*	Length (bp/aa)	Exons	Domains (CDD)	PKS product
37260	7823/2463	9	KS AT DH ER – MT – ACP KR	
39026	7405/2122	12	KR AT DH - ER KR ACP	
41618	11394/3538	14	KS AT DH MT – KR ACP C A ACP	
41846	5565/1855	1	A ACP KS AT ACP	
43495	8213/2314	19	KS AT DH MT ER KR ACP	
44005	7444/2474	2	KS AT DH MT – KR ACP	Asperrubrol (this study)

Continued on Next Page

Table 8.1 – Continued

Model ID*	Length (bp/aa)	Exons	Domains (CDD)	PKS product
44965	5340/1780	1	KS AT DH(TH) KR ACP	6-MSA (this study)
47991	7612/2433	6	KS AT DH MT ER KR ACP	
51499	5419/1718	5	KS AT – ACP	Coumarins ¹
56896	6755/2129	6	AT KS AT – ACP TE	YWA1 (this study)
56946	7915/2600	3	KS AT – ACP TE	Azaphilones ²
118581	11427/3760	5	KS AT DH KR KR ACP C A ACP TR	
118598	8422/2441	15	KS AT DH MT ER KR ACP	
118617	8542/2472	19	KS AT DH – ER KR ACP	
118624	11791/3645	7	KS AT DH MT KR C A ACP TR	
118629	7409/2442	3	KS AT DH – KR ACP	
118644	11948/3706	7	KS AT DH – KR ACP C A ACP TR	
118666	6733/2199	4	KS AT ACP – ER KR ACP	
118744	7394/2347	5	KS AT DH MT ER KR ACP	
128601	11657/3779	3	KS AT DH MT KR C A ACP (TR)	
128638	6984/2202	6	KS AT DH – ER KR ACP	
138585	7869/2421	11	KS AT DH MT ER KR ACP	
171221	8222/2358	10	KS AT DH MT ER KR	
176722	12144/3942	6	KS AT DH MT – ACP C A ACP TR	
179079	8683/2668	10	KS AT DH MT ER KR ACP	
(179462)			SAT KS AT PT ACP	Tetracyclins ³
179585	7090/2347	2	KS AT DH KR ER ACP	
181803	8305/2591	9	KS AT DH MT ER KR ACP	
187099	7099/2211	3	KR AT DH – ER KR ACP	
188697	8178/2469	14	KS AT DH – ER KR ACP	
188817	8248/2534	11	KS AT DH MT ER KR ACP	
190014	7760/2517	4	KS AT ACP MT TR	
191422	6664/2122	7	– KS AT – ACP TE	YWA1 (this study)
191702	8153/2494	13	KS AT DH MT ER KR	
194381	6957/2104	3	KS AT DH ACP MT	
210217	8194/2560	8	KS AT DH MT ER KR ACP	
211885	7996/2443	10	KS AT DH MT KR ACP	
225574	8010/2527	7	KS AT DH MT ER KR ACP	Fumonisin ⁴

8.3 Assessment of *A. nidulans* as a host for heterologous PKS expression

8.3.1 Only three compounds were linked to a PKS

As only three compound were produced (and only two detected by metabolite profile analysis) from heterologous expression of 37 genes, the success rate was indeed low (8.1 %). Although, there are many possible explanations for this result, some are more likely than others.

Were the genes correctly transcribed? The main question here is whether or not the ORFs of the genes were correctly predicted. I have used the automated FGeneSH gene prediction applied on the Joint Genome Institute (JGI) database. With only little knowledge of PKSs, it is quite possible that some of the predictions are wrong. Curiously, this was in fact the case when it was first discovered that YWA1 was the product of WA in *A. nidulans*; Watanabe and co-workers first reported that WA produced the polyketide citreoisocoumarin, which is a bicyclic compound with a C₅ side chain [Watanabe et al., 1998]. Later they realized that they had used a C-terminal truncated version, which lacked part of the TE domain. Expression of the correct ORF resulted in production of YWA1, which is the true polyketide product of WA [Watanabe et al., 1999].

Are the genes correctly translated and functional? All the PKSs, except one, contain introns. There is, to the best of my knowledge, no studies on the evolution of introns or splicing of introns, why it is possible that *A. nidulans* is unable to correctly splice some of the introns from *A. niger* leading to wrong translation of the mRNA. Moreover, even if the mRNA is translated correctly the huge enzyme may not be able to fold in a correct manner. The PKS may depend on co-factors for proper allostery, or on other proteins for protein-protein interactions. This may be with respect to formation of an active enzyme complex, or delivery or reception of the PK compound from another enzyme: The PKS may depend on the delivery of a starter unit before synthesis can occur, or it may be dependent on another enzyme to receive or release the PK compound once the PKS has finished synthesis. When assessing the research that has been done on fungal PKSs within the past five years, it seems as there are only few PKSs that are not dependent on alternative starter units or external release mechanisms. Some of the PKSs that are independent of such comprise mostly 6-MSA synthases and non-reducing PKSs e.g. YWA1 synthases and orsellinic acid synthases. Some gene clusters comprise two synthases which must both be expressed in order to synthesize the final compound, which is the case for e.g. lovastatin and the recently described azanigerones [Zabala et al., 2012]. Furthermore, alternative starter unit are employed by several PKSs: the sterigmatocystin/aflatoxin PKS uses a hexanoyl-CoA starter unit that is produced by a fatty acid synthase embedded in the gene cluster [Crawford et al., 2006], and the asperribrol PKS from *A. niger* is dependent on a cinnamic acid starter unit (this study).

In my personal opinion, the latter option seems the most probable as the synthesis of PK compounds is very complex, and has a high degree of variability. However, I would assume that problems with wrong gene prediction and maybe even wrong intron splicing may be the case for a few of the PKSs.

Based on the fact that there was so many parameters to assess in order to fully understand the outcome of this study, I decided not to submit this study for publication. Rather, I used the results of the study to perform in-depth studies on the PKSs that did provide new compounds. Additionally, I used this semi-high throughput (semi-HTP) study to locate bottlenecks and assess current tools available for strain construction.

8.3.2 There are obvious bottlenecks in HTP strain construction

As briefly mentioned above, this semi-HTP study also served as a test-case for performing HTP strain construction. Doing so, I identified several bottlenecks which preferably should be assessed and resolved before initiating a HTP strain construction study.

Screening of clones could be tedious, and the cloning efficiency varied from batch to batch. Generally, the uracil-excision cloning system is very robust and has a high efficiency, but it does have one disadvantage: digestion of the cassette. Recall that a uracil-excision cassette contains a restriction site flanked by two nicking sites. When digested the vector is linearized by the restriction enzyme, and the ends are processed by the nicking enzyme to form 6-8 nt single-stranded overhangs. If not fully digested, the reaction will contain both correctly digested vector, partially digested vector and intact vector. The two latter variants will result in false positive clones, increasing the work load for screening of clones. One way to address this issue is by elimination of the UE cassette, and merely fuse PCR fragments. This is indeed possible, however, the cloning efficiency decreases as the number of fragments increases. Therefore this approach is not suitable for a HTP study. Another way to address the problem is by inserting a reporter gene or a killer gene into the UE cassette, which will be disrupted if the fragment is inserted into the vector. This was the principle of the old-fashioned pBluescript vectors which contain a *lacZ* gene with a multiple cloning site, resulting in a color change of the colony from blue to white when exposed to X-gal if the gene was inserted into the multiple cloning site.

Screening of fungal transformants is a hassle. The quick and dirty solution for screening fungal transformant is *spore PCR*, the fungal equivalent of colony PCR. However easy, this method is not very robust, and often it does not work. Therefore it may be necessary to extract genomic DNA, and subject it to PCR analysis e.g. in combination with southern blot analysis. This approach works well, and is not considered laborious when constructing a single strain. However, when constructing several strains, this step becomes extremely tedious and time consuming.

I have addressed the above mentioned obstacles in HTP strain construction by developing a background-free cloning vector that contains the *ccdB* suicide gene in between two identical UE cassettes. This allowed for background-free insertion of genes into the vector, and is highly suitable for HTP plasmid construction. Furthermore, I have developed a background-free transformation system for heterologous or ectopic gene expression in *A. nidulans*, which exploits use of a color marker for efficient verification of correct transformants. These tools are described in Chapter 3, and have been applied throughout this study.

8.4 Experimental

8.4.1 Fungal strains

The *A. nidulans* strain NID3 (*argB2 pyrG89 veA1 nkuA-trS*) [Nielsen et al., 2008] was used as a reference and for insertion of all genes. For *A. niger*, KB1001 (*pyrGΔ kusA::AFpyrG*) [Chiang et al., 2011] was used as a reference and for gene deletions. *A. niger* deletion strains comprised a 44005Δ strain (*pyrGΔ kusA::AFpyrG* ASPNIDRAFT_44005Δ::*hph*) and a NRPSΔ strain (*pyrGΔ kusA::AFpyrG* ASPNIDRAFT_190891Δ::*hph*).

Table 8.2: *A. nidulans* strains used and constructed in this study. *ori^R** contains a point mutation eliminating a Nb.BtsI nicking site.

Strain	Genotype	Source
Strain	Genotype	Source
NID3	<i>argB2 pyrG89 veA1 nkuA-trS</i>	[Nielsen et al., 2008]
37260	<i>argB2 pyrG89 veA1 nkuA-trS</i> IS1::P <i>gpdA</i> ::ASPNIDRAFT_37260::T <i>trpC</i> :: <i>argB</i>	This study
39026	<i>argB2 pyrG89 veA1 nkuA-trS</i> IS1::P <i>gpdA</i> ::ASPNIDRAFT_39026::T <i>trpC</i> :: <i>argB</i>	This study
41618	<i>argB2 pyrG89 veA1 nkuA-trS</i> IS1::P <i>gpdA</i> ::ASPNIDRAFT_41618::T <i>trpC</i> :: <i>argB</i>	This study
41846	<i>argB2 pyrG89 veA1 nkuA-trS</i> IS1::P <i>gpdA</i> ::ASPNIDRAFT_41846::T <i>trpC</i> :: <i>argB</i>	This study
43495	<i>argB2 pyrG89 veA1 nkuA-trS</i> IS1::P <i>gpdA</i> ::ASPNIDRAFT_43495::T <i>trpC</i> :: <i>argB</i>	This study
44005	<i>argB2 pyrG89 veA1 nkuA-trS</i> IS1::P <i>gpdA</i> ::ASPNIDRAFT_44005::T <i>trpC</i> :: <i>argB</i>	This study
44005+TE	<i>argB2 pyrG89 veA1</i> IS1::P <i>gpdA</i> ::ASPNIDRAFT_44005::T <i>trpC</i> :: <i>argB</i> pDH121	This study
44965	<i>argB2 pyrG89 veA1 nkuA-trS</i> IS1::P <i>gpdA</i> ::ASPNIDRAFT_44965::T <i>trpC</i> :: <i>argB</i>	This study
47991	<i>argB2 pyrG89 veA1 nkuA-trS</i> IS1::P <i>gpdA</i> ::ASPNIDRAFT_47991::T <i>trpC</i> :: <i>argB</i>	This study
51499	<i>argB2 pyrG89 veA1 nkuA-trS</i> IS1::P <i>gpdA</i> ::ASPNIDRAFT_51499::T <i>trpC</i> :: <i>argB</i>	This study
56896	<i>argB2 pyrG89 veA1 nkuA-trS</i> IS1::P <i>gpdA</i> ::ASPNIDRAFT_56896::T <i>trpC</i> :: <i>argB</i>	This study
56946	<i>argB2 pyrG89 veA1 nkuA-trS</i> IS1::P <i>gpdA</i> ::ASPNIDRAFT_56946::T <i>trpC</i> :: <i>argB</i>	This study
118581	<i>argB2 pyrG89 veA1 nkuA-trS</i> IS1::P <i>gpdA</i> ::ASPNIDRAFT_118581::T <i>trpC</i> :: <i>argB</i>	This study
118598	<i>argB2 pyrG89 veA1 nkuA-trS</i> IS1::P <i>gpdA</i> ::ASPNIDRAFT_118598::T <i>trpC</i> :: <i>argB</i>	This study
118617	<i>argB2 pyrG89 veA1 nkuA-trS</i> IS1::P <i>gpdA</i> ::ASPNIDRAFT_118617::T <i>trpC</i> :: <i>argB</i>	This study
118624	<i>argB2 pyrG89 veA1 nkuA-trS</i> IS1::P <i>gpdA</i> ::ASPNIDRAFT_118624::T <i>trpC</i> :: <i>argB</i>	This study
118629	<i>argB2 pyrG89 veA1 nkuA-trS</i> IS1::P <i>gpdA</i> ::ASPNIDRAFT_118629::T <i>trpC</i> :: <i>argB</i>	This study
118644	<i>argB2 pyrG89 veA1 nkuA-trS</i> IS1::P <i>gpdA</i> ::ASPNIDRAFT_118644::T <i>trpC</i> :: <i>argB</i>	This study
118666	<i>argB2 pyrG89 veA1 nkuA-trS</i> IS1::P <i>gpdA</i> ::ASPNIDRAFT_118666::T <i>trpC</i> :: <i>argB</i>	This study
118744	<i>argB2 pyrG89 veA1 nkuA-trS</i> IS1::P <i>gpdA</i> ::ASPNIDRAFT_118744::T <i>trpC</i> :: <i>argB</i>	This study
128601	<i>argB2 pyrG89 veA1 nkuA-trS</i> IS1::P <i>gpdA</i> ::ASPNIDRAFT_128601::T <i>trpC</i> :: <i>argB</i>	This study

Continued on Next Page

Table 8.2 – Continued

Strain	Genotype	Source
128638	<i>argB2 pyrG89 veA1 nkuA-trS</i> IS1::P <i>gpdA</i> ::ASPNIDRAFT_128638::T <i>trpC</i> :: <i>argB</i>	This study
138585	<i>argB2 pyrG89 veA1 nkuA-trS</i> IS1::P <i>gpdA</i> ::ASPNIDRAFT_138585::T <i>trpC</i> :: <i>argB</i>	This study
171221	<i>argB2 pyrG89 veA1 nkuA-trS</i> IS1::P <i>gpdA</i> ::ASPNIDRAFT_171221::T <i>trpC</i> :: <i>argB</i>	This study
176722	<i>argB2 pyrG89 veA1 nkuA-trS</i> IS1::P <i>gpdA</i> ::ASPNIDRAFT_176722::T <i>trpC</i> :: <i>argB</i>	This study
179079	<i>argB2 pyrG89 veA1 nkuA-trS</i> IS1::P <i>gpdA</i> ::ASPNIDRAFT_179079::T <i>trpC</i> :: <i>argB</i>	This study
179079+TE	<i>argB2 pyrG89 veA1</i> IS1::P <i>gpdA</i> ::ASPNIDRAFT_179079::T <i>trpC</i> :: <i>argB</i> pDH123	This study
179585	<i>argB2 pyrG89 veA1 nkuA-trS</i> IS1::P <i>gpdA</i> ::ASPNIDRAFT_179585::T <i>trpC</i> :: <i>argB</i>	This study
179585+TE	<i>argB2 pyrG89 veA1</i> IS1::P <i>gpdA</i> ::ASPNIDRAFT_179585::T <i>trpC</i> :: <i>argB</i> pDH122	This study
181803	<i>argB2 pyrG89 veA1 nkuA-trS</i> IS1::P <i>gpdA</i> ::ASPNIDRAFT_181803::T <i>trpC</i> :: <i>argB</i>	This study
187099	<i>argB2 pyrG89 veA1 nkuA-trS</i> IS1::P <i>gpdA</i> ::ASPNIDRAFT_187099::T <i>trpC</i> :: <i>argB</i>	This study
188697	<i>argB2 pyrG89 veA1 nkuA-trS</i> IS1::P <i>gpdA</i> ::ASPNIDRAFT_188697::T <i>trpC</i> :: <i>argB</i>	This study
188817	<i>argB2 pyrG89 veA1 nkuA-trS</i> IS1::P <i>gpdA</i> ::ASPNIDRAFT_188817::T <i>trpC</i> :: <i>argB</i>	This study
190014	<i>argB2 pyrG89 veA1 nkuA-trS</i> IS1::P <i>gpdA</i> ::ASPNIDRAFT_190014::T <i>trpC</i> :: <i>argB</i>	This study
191422	<i>argB2 pyrG89 veA1 nkuA-trS</i> IS1::P <i>gpdA</i> ::ASPNIDRAFT_191422::T <i>trpC</i> :: <i>argB</i>	This study
191702	<i>argB2 pyrG89 veA1 nkuA-trS</i> IS1::P <i>gpdA</i> ::ASPNIDRAFT_191702::T <i>trpC</i> :: <i>argB</i>	This study
194381	<i>argB2 pyrG89 veA1 nkuA-trS</i> IS1::P <i>gpdA</i> ::ASPNIDRAFT_194381::T <i>trpC</i> :: <i>argB</i>	This study
210217	<i>argB2 pyrG89 veA1 nkuA-trS</i> IS1::P <i>gpdA</i> ::ASPNIDRAFT_210217::T <i>trpC</i> :: <i>argB</i>	This study
211885	<i>argB2 pyrG89 veA1 nkuA-trS</i> IS1::P <i>gpdA</i> ::ASPNIDRAFT_211885::T <i>trpC</i> :: <i>argB</i>	This study
225574	<i>argB2 pyrG89 veA1 nkuA-trS</i> IS1::P <i>gpdA</i> ::ASPNIDRAFT_225574::T <i>trpC</i> :: <i>argB</i>	This study

8.4.2 Media

Minimal medium for *A. nidulans* was made as described in [Cove, 1966] but with 1% glucose, 10 mM NaNO₃ and 2% agar. Minimal medium for *A. niger* was prepared as described in [Chiang et al., 2011]. YES, MEA and CYA media were prepared as described by Frisvad and Samson [Samson et al.]. When necessary, media were supplemented with 4 mM L-arginine, 10 mM uridine, 10 mM uracil, and/or 100 μ g/ml Hygromycin B (Invivogen, San Diego, CA, USA).

8.4.3 PCR, cloning and transformation

Primers for amplification of genes were designed using the Primer Help for USER software (PHUSER) [Olsen et al., 2011]. All primers are listed in Table 8.4.3. All PCR reactions were carried out using the *PfuX7* polymerase [Nørholm, 2010], and either HF or Phire PCR buffer (Finnzymes/Thermo Scientific). The PCR program was: 98°C for 1 min, 98°C for 15 s, 60°C for 30 s, 72°C for 30s/kb, and 72°C for 5 min. Normally,

I performed 35 cycles of amplification. Diagnostic (conidial) PCR was performed as a regular PCR reaction, except initial denaturation step was 20 min.

ASPNIDRAFT_44965 was amplified with BGHA582-585 (2 fragments), 56896 with BGHA586-591 (3 fragments), 128601 with BGHA592-607 (5 fragments), 194381 with BGHA608-613 (3 fragments), 211885 with BGHA614-621 (4 fragments), and 225574 with BGHA622-629 (4 fragments). 41618 (primers 47-50), 118581 (primers 63-66), 118624 (primers 71-74), 118644 (primers 77-80), and 176722 (primers 91-94) were amplified as two fragments, otherwise the genes were amplified as a single fragment (Table 8.4.3 lists the pairwise primer sets).

Fragments were assembled via uracil-excision fusion [Geu-Flores et al., 2007] into a compatible vector, and correct clones were verified by restriction digestion (plasmids are listed in Table 8.4.3). Protoplasting and gene-targeting procedures were performed as described previously [Johnstone et al., 1985, Nielsen et al., 2006].

Table 8.3: Primers used in this study. All primers are shown in 5'-3' direction

Primer number	Primer name	Sequence
1	BGHA582	AGAGCGAUATGCCAGGCCTTGTACAC
2	BGHA583	ACGGGGTCCUCGGTACATTCCCCATAGTTT
3	BGHA584	AGGACCCCGUTCTCGTCACCCCTACCAT
4	BGHA585	TCTGCGAUTTAAGCATCCAGCTCCTTTGT
5	BGHA586	AGAGCGAUATGGAGGGTCCATCTCGTGTGT
6	BGHA587	AGAACCGAUATGAAGCGATTGACCGGGAC
7	BGHA588	ATCGGTTTCUGCCAAGGCAAACATTGGAC
8	BGHA589	AAAGAAUCGGACAGTGCAAGTTGCG
9	BGHA590	ATTCTTUGACTGCGCTGCTGCGGAG
10	BGHA591	TCTGCGAUTTAATTGGCCATGGCGTTGCCA
11	BGHA592	AGAGCGAUATGTCTTCTTCACCCAAAGAGCCC
12	BGHA593	AGTGGCGCUAAAGAAGCCGGCGGCGTA
13	BGHA594	AGCGCCACUGATGCCATCCGCATTGCCTTCT
14	BGHA595	AGTGTCUCCAGGAGAGGCTTGGTAGCG
15	BGHA596	AGACACUCGGTGCCTTGTA CTCCGCCTACAC
16	BGHA597	ACCGTCGAUCTTGGGGCCCAGAGCCTCGTTCA
17	BGHA598	ATCGACGGUACCATCTACCTGGACGAAGTGT
18	BGHA599	ACCGCAAGUGGCCGCAAGACGGGGCAG
19	BGHA606	ACTTGCGGUGTTGTGCGAGCCCAGCTCC

Continued on Next Page

Table 8.3 – Continued

Primer number	Primer name	Sequence
20	BGHA607	TCTGCGAUCCGGCAATCTTGGCCTTGC
21	BGHA608	AGAGCGAUATGGCAATCATGGAACAG
22	BGHA609	ACCGGCGCUGTTATTGACGTTCCGTACT
23	BGHA610	AGCGCCGGUGTCCAAGTCTCAGCAATCCCTCTA
24	BGHA611	ATCTTGCAUGATTCCGGCAGCGGCCAC
25	BGHA612	ATGCAAGAUTCTGTATACGATCCTCTA
26	BGHA613	TCTGCGAUTTATTATCAAGCGGAGCT
27	BGHA614	AGAGCGAUATGAGTCCCAGCAAGAAT
28	BGHA615	ACGCCGAGCAAUGAAGCAGTTCCATGCAGT
29	BGHA616	ATTGCTCGGCGUTTTACCCGGCCAGGGAGC
30	BGHA617	ATAGCCGCUCAGCATTCTTGAGCAAAGCCTGT
31	BGHA618	AGCGGCTAUTCCATACTGTCAATGACGG
32	BGHA619	AGCATTUGTAATACCCACTGGACTGT
33	BGHA620	AAATGCUCCCACCCACGGGGACTGT
34	BGHA621	TCTGCGAUTCAGACCACCTCCTGTGTACTGAT
35	BGHA622	AGAGCGAUATGGATAGTCAAGACTCGTCTCCC
36	BGHA623	AGCTTTCTUTTAGGGATGCCGGGTGTGC
37	BGHA624	AAGAAAGCUTAGTCGAGCATGAACGATACATCA
38	BGHA625	AGAATCAAUATGCGAGCTATCGCTGGC
39	BGHA626	ATTGATTCUCTTGAGCTCGAGCTCGTC
40	BGHA627	ATCGTGGCCGCUTCTTCGTATGAGAGGAACGATG
41	BGHA628	AGCGGCCACGAUGCCATGTGTTTACGCTACT
42	BGHA629	TCTGCGAUTCACGGGGCCTTCATTTC
43	FW-37260	AGAGCGAUATGTCTGCCAGCGAGACT
44	RV-37260	TCTGCGAUTCACTCCTTCACCCCATATCT
45	FW-39026	AGAGCGAUATGGAAGAGCACGTCAGAGA
46	RV-39026	TCTGCGAUCGGCTGTGCCCCGATAAAC
47	FW-41618-f1	AGAGCGAUATGAACGTCCCCAATGAG
48	RV-41618-f1	ATCAACCUCATGCACTATAGACCGGC
49	FW-41618-f2	AGGTTGAUCCGTCCACAGCGGCCCTA
50	RV-41618-f2	TCTGCGAUCTAGGCATGAGCCCCTTTGCTTTC
51	FW-41846	AGAGCGAUATGGTACCAGAGGACGACCAA
52	RV-41846	TCTGCGAUCTACTGACCTTCTACCTGCTTCT

Continued on Next Page

Table 8.3 – Continued

Primer number	Primer name	Sequence
53	FW-43495	AGAGCGAUATGACTACTCCGATGAACG
54	RV-43495	TCTGCGAUTCACTCACCACATTACC
55	FW-44005	AGAGCGAUATGACACGAACGGAGCCAAT
56	RV-44005	TCTGCGAUTTATGCCTGCAGCTGCGG
57	FW-47991	AGAGCGAUATGGCCCTTCGTCTTCCC
58	RV-47991	TCTGCGAUTCACTCGATCCCCAACAAC
59	FW-51499	AGAGCGAUATGAGGCTCTTCTTCTTTAGTA
60	RV-51499	TCTGCGAUTTACCAAGACGCAGTGAG
61	FW-56946	AGAGCGAUATGGTCACAACGACAAGG
62	RV-56946	TCTGCGAUTCAAGGATTCAAAAATCCCATCT
63	FW-118581-1	AGAGCGAUATGGTTTCTGAATATGTCCC
64	RV-118581-1	ACCGCAGACUGGAGTGATGACTCAAACG
65	FW-118581-2	AGTCTGCGGUTCGGCATGTTGAGTCTTT
66	RV-118581-2	TCTGCGAUCTAAACTAGTAATGACCTCAGACT
67	FW-118598	AGAGCGAUATGTCTTTGAAGTTTCCGGAAGA
68	RV-118598	TCTGCGAUTCATCGTCGGTATTCGCTCTT
69	FW-118617	AGAGCGAUATGGAAGGGAAGCATTTAGAACCC
70	RV-118617	TCTGCGAUTCAGGACCAACCACGAGG
71	FW-118624-f1	AGAGCGAUAATATGGCATACCCACAAG
72	RV-118624-f1	AAGATGCAUCGTCGTCAACCATCAAGG
73	FW-118624-f2	ATGCATCTUACTTTCTCTGCTTGGGGAAGGC
74	RV-118624-f2	TCTGCGAUCGTGGAACGAGTGGTGGTGAT
75	FW-118629	AGAGCGAUATGTCAGTCCCTGAACCC
76	RV-118629	TCTGCGAUTAGCCAGCTCTTGGAATTG
77	FW-118644-f1	AGAGCGAUGACATCATGGCAACGGAACGA
78	RV-118644-f1	ATTCTGUGGCTTTGAAGGAGGCCTAAGTC
79	FW-118644-f2	ACAGAAUCGGTGTCCTGAATTATT
80	RV-118644-f2	TCTGCGAUTTCCAAATAAGCAGCAAC
81	FW-118666	AGAGCGAUATGAATAGGTCACAGTCTGCG
82	RV-118666	TCTGCGAUTTACATCCCCGACTTACC
83	FW-118744	AGAGCGAUATAGGCTGTCGCCTACCG
84	RV-118744	TCTGCGAUCATCCTCGCATCCATCAACT
85	FW-128638	AGAGCGAUCCGATTGCAATTGTAGGCATGT

Continued on Next Page

Table 8.3 – Continued

Primer number	Primer name	Sequence
86	RV-128638	TCTGCGAUCTCCAAACTCGACTTCTTCAAC
87	FW-138585	AGAGCGAUATGGGAATGAGACTCCCT
88	RV-138585	TCTGCGAUGCGTAAGCTCATCCAGCT
89	FW-171221	AGAGCGAUATGTCTACACGTCCCCAAT
90	RV-171221	TCTGCGAUCTGCTCACCCCTCCTTCAC
91	FW-176722-f1	AGAGCGAUATGGCCCCTCATCCACAA
92	RV-176722-f1	AATTGAUGCGGCGAGAAGGTGATTGG
93	FW-176722-f2	ATCAATUCTGGCTTCCATCCCCGCC
94	RV-176722-f2	TCTGCGAUTCACCGTCGGGATTCCAACC
95	FW-179079	AGAGCGAUATGGCTTTCGACAATTTCGCACAA
96	RV-179079	TCTGCGAUTCAAACCCACCCCTTCTTCAAC
97	FW-179585	AGAGCGAUATGGACACACCCCTCTCA
98	RV-179585	TCTGCGAUGTCGTTGCTTGTGGTGATC
99	FW-181803	AGAGCGAUATGGAGGACCCTACTACT
100	RV-181803	TCTGCGAUCTAATTCTGTGACTGAAACTG
101	FW-187099	AGAGCGAUATGGCTGGAGACTTAAGGG
102	RV-187099	TCTGCGAUCACAGCCTTCTCAACCAA
103	FW-188697	AGAGCGAUATGAGTAACAATGAGGAGC
104	RV-188697	TCTGCGAUTCACTGCTTGCTATACTT
105	FW-188817	AGAGCGAUATGGCCGGAGAGTATTCAACG
106	RV-188817	TCTGCGAUCTACGTCCGCTTCGTTGC
107	FW-190014	AGAGCGAUATGCCCGCCAATACTCCA
108	RV-190014	TCTGCGAUCAGAAAACCAGTTTGCCTCC
109	FW-191422	AGAGCGAUATGGAGACAGTCGCCGAGGTA
110	RV-191422	TCTGCGAUCTACAACAAAGCCCGGCGAAGGA
111	FW-191702	AGAGCGAUATGGCCATAAACAATGGA
112	RV-191702	TCTGCGAUCTCAGCACTAATGCCATA
113	FW-210217	AGAGCGAUATGCAACTACTAACGACT
114	RV-210217	TCTGCGAUTTACAAATGACAGGTATTCT
115	44004-fw	AGAGCGAUATGGACATTGGTGTAGGCGTC
116	44004-rv	TCTGCGAUTCATTCGGGATTTGGCTGTT
117	142878-fw	AGAGCGAUATGGCCTCTTCCACTGTTGTC
118	142878-rv	TCTGCGAUTCACGGGGCATGCTTAAAAAAG

Continued on Next Page

Table 8.3 – Continued

Primer number	Primer name	Sequence
119	179510-fw	AGAGCGAUATGTCCAGCGAAAATTTCACTATAAC
120	179510-rv	TCTGCGAUTTACAGCTTGTCTTCCTGGCT
121	50432-bi1-fw	AGAGCGAUATGCTTGATAAACACATACCAGATG
122	50432-bi1-rv	TCTGCGAUTCATTGTCGAAAAGAATTAAGGATT
123	179510-rv	TCTGCGAUTTACAGCTTGTCTTCCTGGCT
124	179510-rv	TCTGCGAUTTACAGCTTGTCTTCCTGGCT
125	179510-rv	TCTGCGAUTTACAGCTTGTCTTCCTGGCT
126	179510-rv	TCTGCGAUTTACAGCTTGTCTTCCTGGCT

Table 8.4: Plasmids used and constructed in this study. All plasmids are *ori^R amp^R*. *hph^{*}* contains a silent point mutation eliminating an AsiSI restriction site. *ori^R** contains a point mutation eliminating a Nb.BtsI nicking site. UEC = AsiSI/Nb.BtsI uracil excision cassette

Plasmid	Genotype	Source
pCMBU1111	TS1::P <i>gpdA</i> ::UEC::T <i>trpC</i> :: <i>argB</i> ::TS2	[Hansen et al., 2011]
pDHX2	AMA1 P <i>gpdA</i> _{0.5kb} ::UEC::T <i>trpC</i> :: <i>pyrG</i>	This study
pCMBU0020	UEC	[Hansen et al., 2011]
pCMBU0020	US-44005:: <i>hph</i> ::DS-44005	[Hansen et al., 2011]
pCMBU0020	US-190891:: <i>hph</i> ::DS-190891	[Hansen et al., 2011]
pDH15	TS1::P <i>gpdA</i> ::ASPNIDRAFT_37260::T <i>trpC</i> :: <i>argB</i> ::TS2	This study
pDH16	TS1::P <i>gpdA</i> ::ASPNIDRAFT_39026::T <i>trpC</i> :: <i>argB</i> ::TS2	This study
pDH17	TS1::P <i>gpdA</i> ::ASPNIDRAFT_41618::T <i>trpC</i> :: <i>argB</i> ::TS2	This study
pDH18	TS1::P <i>gpdA</i> ::ASPNIDRAFT_41846::T <i>trpC</i> :: <i>argB</i> ::TS2	This study
pDH19	TS1::P <i>gpdA</i> ::ASPNIDRAFT_43495::T <i>trpC</i> :: <i>argB</i> ::TS2	This study
pDH20	TS1::P <i>gpdA</i> ::ASPNIDRAFT_44005::T <i>trpC</i> :: <i>argB</i> ::TS2	This study
pDH25	TS1::P <i>gpdA</i> ::ASPNIDRAFT_44965::T <i>trpC</i> :: <i>argB</i> ::TS2	This study
pDH26	TS1::P <i>gpdA</i> ::ASPNIDRAFT_47991::T <i>trpC</i> :: <i>argB</i> ::TS2	This study
pDH27	TS1::P <i>gpdA</i> ::ASPNIDRAFT_51499::T <i>trpC</i> :: <i>argB</i> ::TS2	This study
pDH28	TS1::P <i>gpdA</i> ::ASPNIDRAFT_56896::T <i>trpC</i> :: <i>argB</i> ::TS2	This study
pDH29	TS1::P <i>gpdA</i> ::ASPNIDRAFT_56946::T <i>trpC</i> :: <i>argB</i> ::TS2	This study
pDH30	TS1::P <i>gpdA</i> ::ASPNIDRAFT_118581::T <i>trpC</i> :: <i>argB</i> ::TS2	This study
pDH31	TS1::P <i>gpdA</i> ::ASPNIDRAFT_118598::T <i>trpC</i> :: <i>argB</i> ::TS2	This study

Continued on Next Page

Table 8.4 – Continued

Plasmid	Genotype	Source
pDH32	TS1::P <i>gpdA</i> ::ASPNI DRAFT_118617::T <i>trpC</i> :: <i>argB</i> ::TS2	This study
pDH33	TS1::P <i>gpdA</i> ::ASPNI DRAFT_118624::T <i>trpC</i> :: <i>argB</i> ::TS2	This study
pDH34	TS1::P <i>gpdA</i> ::ASPNI DRAFT_118629::T <i>trpC</i> :: <i>argB</i> ::TS2	This study
pDH35	TS1::P <i>gpdA</i> ::ASPNI DRAFT_118644::T <i>trpC</i> :: <i>argB</i> ::TS2	This study
pDH36	TS1::P <i>gpdA</i> ::ASPNI DRAFT_118666::T <i>trpC</i> :: <i>argB</i> ::TS2	This study
pDH37	TS1::P <i>gpdA</i> ::ASPNI DRAFT_118744::T <i>trpC</i> :: <i>argB</i> ::TS2	This study
pDH38	TS1::P <i>gpdA</i> ::ASPNI DRAFT_128601::T <i>trpC</i> :: <i>argB</i> ::TS2	This study
pDH39	TS1::P <i>gpdA</i> ::ASPNI DRAFT_128638::T <i>trpC</i> :: <i>argB</i> ::TS2	This study
pDH40	TS1::P <i>gpdA</i> ::ASPNI DRAFT_138585::T <i>trpC</i> :: <i>argB</i> ::TS2	This study
pDH41	TS1::P <i>gpdA</i> ::ASPNI DRAFT_171221::T <i>trpC</i> :: <i>argB</i> ::TS2	This study
pDH42	TS1::P <i>gpdA</i> ::ASPNI DRAFT_176722::T <i>trpC</i> :: <i>argB</i> ::TS2	This study
pDH43	TS1::P <i>gpdA</i> ::ASPNI DRAFT_179079::T <i>trpC</i> :: <i>argB</i> ::TS2	This study
pDH44	TS1::P <i>gpdA</i> ::ASPNI DRAFT_179585::T <i>trpC</i> :: <i>argB</i> ::TS2	This study
pDH45	TS1::P <i>gpdA</i> ::ASPNI DRAFT_181803::T <i>trpC</i> :: <i>argB</i> ::TS2	This study
pDH46	TS1::P <i>gpdA</i> ::ASPNI DRAFT_187099::T <i>trpC</i> :: <i>argB</i> ::TS2	This study
pDH47	TS1::P <i>gpdA</i> ::ASPNI DRAFT_188697::T <i>trpC</i> :: <i>argB</i> ::TS2	This study
pDH48	TS1::P <i>gpdA</i> ::ASPNI DRAFT_188817::T <i>trpC</i> :: <i>argB</i> ::TS2	This study
pDH49	TS1::P <i>gpdA</i> ::ASPNI DRAFT_190014::T <i>trpC</i> :: <i>argB</i> ::TS2	This study
pDH50	TS1::P <i>gpdA</i> ::ASPNI DRAFT_191422::T <i>trpC</i> :: <i>argB</i> ::TS2	This study
pDH51	TS1::P <i>gpdA</i> ::ASPNI DRAFT_191702::T <i>trpC</i> :: <i>argB</i> ::TS2	This study
pDH52	TS1::P <i>gpdA</i> ::ASPNI DRAFT_194381::T <i>trpC</i> :: <i>argB</i> ::TS2	This study
pDH53	TS1::P <i>gpdA</i> ::ASPNI DRAFT_210217::T <i>trpC</i> :: <i>argB</i> ::TS2	This study
pDH54	TS1::P <i>gpdA</i> ::ASPNI DRAFT_211885::T <i>trpC</i> :: <i>argB</i> ::TS2	This study
pDH55	TS1::P <i>gpdA</i> ::ASPNI DRAFT_225574::T <i>trpC</i> :: <i>argB</i> ::TS2	This study
pDH121	AMA1 P <i>gpdA</i> _{0.5kb} ::44004::T <i>trpC</i> :: <i>pyrG</i>	This study
pDH122	AMA1 P <i>gpdA</i> _{0.5kb} ::142878::T <i>trpC</i> :: <i>pyrG</i>	This study
pDH123	AMA1 P <i>gpdA</i> _{0.5kb} ::179510::T <i>trpC</i> :: <i>pyrG</i>	This study
pDH124	TS1::P <i>gpdA</i> ::ASPNI DRAFT_50432::T <i>trpC</i> :: <i>argB</i> ::TS2	This study

8.4.4 Chemical analysis of strains

Metabolites were extracted in small-scale as described in [Smedsgaard, 1997]. Analytical LC-DAD-MS experiments were performed on an Agilent HP 1100 liquid chromatograph with auto sampler, with a DAD

detector system (Wald-brom, Germany) coupled to a liquid chromatography time-of-flight (LCT) oaTOF mass spectrometer (Micromass, Manchester, UK) with a Z-spray source and a Lock Spray probe. 1 μ l of the extracts were injected on a Luna II (C18 (50 x 2 mm, 3 μ m), Phenomenex, Torrance, CA, USA) column. The analysis was performed with a flow rate of 0.3 ml/min with an acetonitrile:water gradient starting at 15:85 and ending at 100:0 in 20 min, hereafter maintaining 100 % acetonitrile for 5 minutes and then returning to the start conditions in 8 minutes. 45 ppm formic acid was added to the solvents before use. The DAD collected UV/VIS spectra from 200-700 nm with a resolution of 4 nm.

Chapter 9

Identification of the naphtho- γ -pyrone gene cluster in *Aspergillus niger*

This chapter contains a manuscript in preparation, intended for submission to Fungal Genetics and Biology, of which I am first author.

In this study I have investigated the polyketide product of the *A. niger* PKS encoded by 191422, see details in Chapter 8. The development of an expression system for *A. niger* made possible this investigation, that resulted in the discovery that *A. niger* possesses two PKSs that produce WYA1, which is the precursor to naphtho- γ -pyrones and black melanin. The gene cluster defined by 191422 contains all activities for production of naphtho- γ -pyrones, although the production of naphtho- γ -pyrones was previously solely assigned to AlbA [Jørgensen et al., 2011, Chiang et al., 2011].

I have carried out all experimental work in this study, including conception of the idea based on results from Chapter 8, experimental design, interpretation of results and writing of manuscript. To complete this study I need to construct two additional strains: a 191422 cluster deletion strain, and a 191422 Δ *albA* Δ double deletion strain. Hopefully, the cluster deletion strain will verify that AlbA is dependent on this gene cluster for production of naphtho- γ -pyrones. The 191422 Δ *albA* Δ double deletion strain will be used to investigate whether or not the "fawn", and not completely white, appearance of the *albA* Δ strain is caused by very little production of YWA by 191422. Supporting information is included in Appendix D.

Identification of the naphtho- γ -pyrone gene cluster in *Aspergillus niger*

Dorte K. Holm¹, Rasmus J. N. Frandsen¹, Uffe H. Mortensen¹, Kristian F. Nielsen^{2,*}

Affiliation:

¹Eukaryotic Molecular Cell Biology group, Center for Microbial Biotechnology, Department of Systems Biology, Soltofts Plads, Technical University of Denmark, 2800 Kgs. Lyngby, Denmark.

²Metabolic Signaling and Regulation group, Center for Microbial Biotechnology, Department of Systems Biology, Soltofts Plads, Technical University of Denmark, 2800 Kgs. Lyngby, Denmark.

***Corresponding author:**

Kristian F. Nielsen, Technical University of Denmark, Department of Systems Biology
Center for Microbial Biotechnology, Metabolic Signaling and Regulation, Building 221
2800 Kgs. Lyngby, Denmark

Email: kfn@bio.dtu.dk

Phone: +45 45252602

Fax: +45 45884922

Abstract

Aspergillus niger is a filamentous fungus of great importance with respect to enzyme and secondary metabolite production. Since the genome sequence of *A. niger* was released, several secondary metabolites have been linked to a gene, including naphtho- γ -pyrones, which were linked to *albA/fwnA*. In this study we have shown that a silent copy of *albA/fwnA* exists in *A. niger*, and that this gene is embedded within a typical secondary metabolite gene cluster, containing all necessary activities for formation of naphtho- γ -pyrones. We have shown that both AlbA/FwnA and the homologue, which we have named NapA, both produce YWA1, however NapA does not seem to be involved in 1,8-dihydroxynaphthalene (DHN) melanin formation. We therefore propose that these two genes were originally intended for separate and compartmentalized processes, but the high activity of AlbA/FwnA may have caused a shift in this regulation. This result is intriguing as it provides an insight into evolution of secondary metabolism, and may form a basis for studying evolution within black *Aspergilli*.

Keywords

Secondary metabolism, naphtho- γ -pyrone, polyketide, biosynthesis, genomics, *Aspergillus niger*.

1. Introduction

Aspergillus niger is an important filamentous fungus, widely used for production of organic acids and enzymes. *A. niger* is able to produce a wide array of secondary metabolites, reviewed in (Nielsen et al., 2009), some of which are harmful to human health e.g. the carcinogenic fumonisins. Recently, Fisch and co-workers performed genomic mining on the full-genome sequenced *A. niger* ATCC1015, and it revealed no less than 33 putative polyketide synthase (PKS) clusters, 15 non-ribosomal peptide synthase (NRPS) clusters and 9 hybrid PKS/NRPS clusters (Fisch et al., 2009). This number greatly exceeds the number of known metabolites from *A. niger*, why it is hypothesized that the remaining gene clusters are non-functional or silenced under standard laboratory conditions. Many different approaches have been successfully carried out in order to scrutinize the biosynthetic potential of filamentous fungi by e.g. expression of local transcription factors (Bergmann et al., 2007, Chiang et al., 2009), targeting general sumylation and histone methylation mechanisms (Szewczyk et al., 2008, Bok et al., 2009, Chiang et al., 2010), and co-cultivation with other organisms (Schroeckh et al., 2009).

In *A. niger* the AlbA/FwnA PKS (encoded by *albA/fwnA*) is required for production of naphtho- γ -pyrones and in addition, AlbA/FwnA is required for production of the conidial black pigment which is characteristic for the group of black Aspergilli (Chiang et al., 2011; Jørgensen et al., 2011a). Neither the biosynthesis of naphtho- γ -pyrones nor DHN- melanin has been shown for *A. niger*, however it has been shown for other Aspergilli that DHN-melanin can be produced from a YWA1 precursor via chain shortening and subsequent polymerization, reviewed by Langfelder and co-workers (2003).

A. niger contains a PKS (termed PKS44 in (Fisch et al., 2009)) that have a high sequence similarity to AlbA, but the function of this PKS has not been reported. Another black *Aspergillus*, *A. carbonarius*, contains homologues of both AlbA and PKS44.

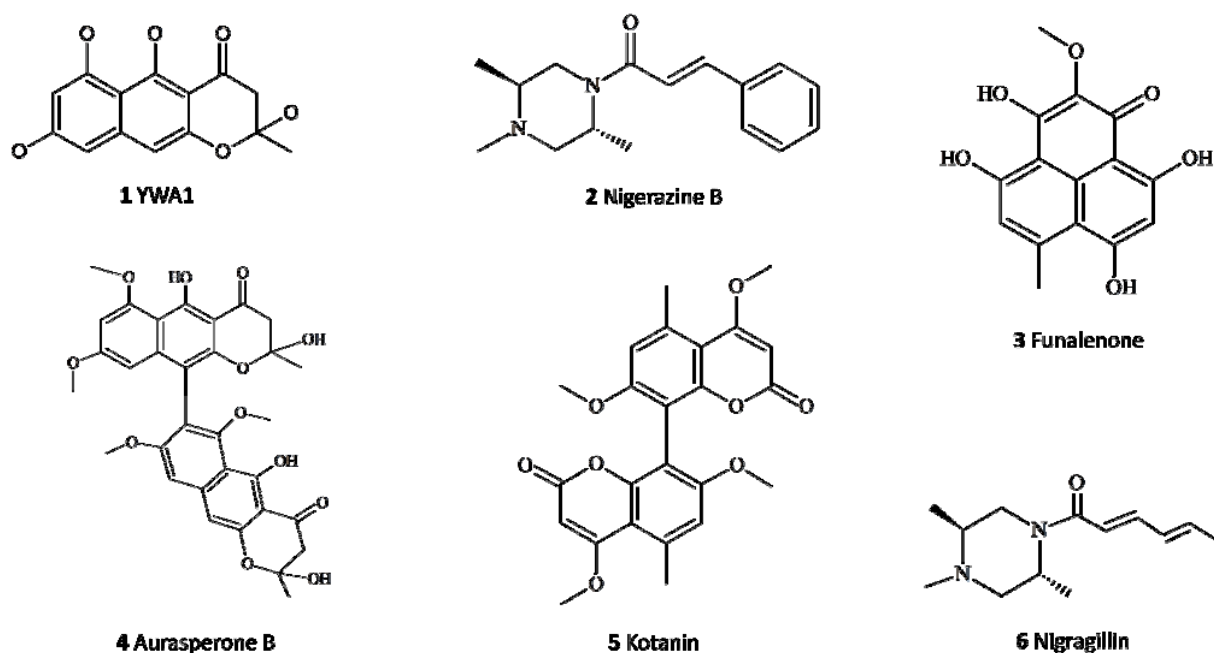


Figure 1. Selected naphtho- γ -pyrone structures from *A. niger*.

2. Materials and methods

2.1. Strains and media

The strain IBT 29539 (*argB2*, *pyrG89*, *veA1*, *nkuAΔ*) was used for strain constructions in *A. nidulans*. ATCC1015-derived strain KB1001 (*pyrGΔ kusaA::AFpyrG*) was used for all strain construction in *A. niger*. All fungal strains constructed in the present work (Appendix, Table A1) have been deposited in the IBT Culture Collection at Department of Systems Biology, DTU, Kgs. Lyngby, Denmark. *Escherichia coli* strain DH5 α was used for propagating all plasmids, except *E. coli* *ccdB* survival2 cells (Invitrogen) were used to propagate plasmids carrying the *ccdB* gene. Minimal medium for *A. nidulans* was made as described in (Cove, 1966) but with 1% glucose, 10 mM NaNO₃ and 2% agar. Minimal medium for *A. niger* was prepared as described in (Chiang *et al.*, 2011). YES media were prepared as described by Frisvad and Samson (Frisvad and Samson, 2004). When necessary, media were supplemented with 4 mM L-arginine, 10mM uridine, 10mM uracil and/or 100 μ g/ml Hygromycin B (Invivogen, San Diego, CA, USA).

2.2. PCR, cloning and transformation

Primers for amplification of genes were designed using the Primer Help for USER software (PHUSER) (Olsen *et al.*, 2011). All primers are listed in Appendix (Table A2). The PfuX7 polymerase (Nørholm, 2010) was used in all PCR reactions. Fragments were assembled via uracil-excision fusion (Geu-Flores *et al.*, 2007) into a compatible vector. Correct clones were verified by restriction digestion (data not shown). Protoplasting and gene-targeting procedures were performed as described previously (Johnstone *et al.*, 1985; Nielsen *et al.*, 2006). Diagnostic (conidial) PCR was performed as a regular PCR reaction, except initial denaturation step was 20 min.

2.3. Construction of the pDHU1 and pDHU2 vectors for integration into *A. niger* IS1 and IS2

The pDHU1 and pDHU2 vectors were constructed by USER fusion of 5 fragments 1) *E. coli* origin of replication (*ori^R*) and the *E. coli* ampicillin resistance gene (*amp^R*) carrying a silent point mutation omitting a Nb.BtsI nicking site, 2) upstream targeting sequence, 3) a PCR fragment containing 0.5 kb of the *PgpdA* promoter, an AsiSI/Nb.BtsI USER cassette containing *ccdB* and *cm^R*, and the TtrpC terminator, 4) downstream targeting sequence 5) selection marker: *E. coli hph*. Fragments 1 and 3 were amplified from pDH57 (Holm *et al.*, in review) (primers 1+2 and primers 3+4), fragment 2 and 4 were amplified from KB1001 genomic DNA (primers 5+6 and 7+8 for pDHU1, primers 11+12 and 13+14 for pDHU2), fragment 5 was amplified from pCB1003 (primers 17+18) (McCluskey *et al.*, 2010). Fragments were assembled as described in (Geu-Flores *et al.*, 2007), using equal concentrations of purified PCR product, and correct clones were verified by restriction digestion (data not shown).

2.4. Heterologous expression of PKSs from *A. nidulans* IS1

The PKS genes PKS44 (ASPNIDRAFT_191422) AlbA (ASPNIDRAFT_56896), and Acar44 (JGI protein ID: 56260) were amplified by PCR (primers 19+20, 21+22, and 23+24, respectively) and inserted into

pDH57 (argB marker). Transformation substrate was excised by NotI digestion and transformed into IBT 29539, and transformants were selected on medium without arginine. Transformants were verified by diagnostic PCR according to (Hansen *et al.*, 2011).

2.5. Construction of deletion strains

All deletions in *A. niger* was done in KB1001: upstream (US) and downstream (DS) fragment were generated by PCR from KB1001 genomic DNA and inserted into the CMBU0020 vector (Hansen *et al.*, 2011) along with the *hph* marker. US and DS fragments for PKS44 Δ were generated using primers 25-28. *hph* was generated using primers 17+18. After insertion of targeting sequences and *hph* into CMBU0020, bipartite transformation substrates were amplified by PCR, using primers 191422_US-FW+29 and 30+191422_DS-RV. Deletion strains were selected on 100 μ g/ml Hygromycin B (Invivogen),

2.6. Ectopic expression of *napA* in *A. niger*

albA (ASPNIIDRAFT_191422) was amplified by PCR (primers 19+20) from KB1001 genomic DNA and inserted into pDHU1 or pDHU2. Transformation substrate was excised by NotI digestion and transformed into KB1001 and transformants were selected on medium with 100 μ g/ml hygromycin B (Invivogen), and verified by diagnostic PCR.

2.7. Strain construction by sexual crossing

Complementation of WA in *A. nidulans* was investigated by sexual crossing. This was performed by plating each parental strain in a line, <5 mm from each other, on plate containing solid MM plus supplements to

complement all auxotrophies. After three days cultivation at 37 °C, a 5x5 mm square (containing both strains) were transferred to MM without supplements. Since the two haploid strains used in a cross contained complementary marker gene combinations, none of the parental strains are able to grow. After 2-3 days incubation at 37 °C, in order to induce meiosis, the plates were sealed with two layers of Parafilm to omit oxygen supply, and wrapped in aluminum foil to omit light, and incubated for two weeks. Plates were then examined for number and size of formed cleistothecia. Ascospores were plated on MM, and correct progeny was verified by diagnostic PCR.

2.8. Chemical analysis of strains

Unless otherwise stated strains were inoculated as a three point inoculation on solid MM media and incubated at 37 °C for 5 days. Extraction of metabolites was performed as described in (Smedsgaard, 1997). 6-MSA was purchased from (Apin Chemicals, Oxon, UK). Analysis was performed using ultra-high-performance liquid chromatography (UHPLC) UV/Vis diode array detector (DAD) high-resolution MS (TOFMS) on a maXis G3 orthogonal acceleration (OA) hexapole–quadrupole time of flight (hQ- TOF) mass spectrometer (Bruker Daltonics, Bremen) equipped with an electrospray injection (ESI) source and connected to an Ultimate 3000 UHPLC system (Dionex, Sunnyvale). The column used was a reverse-phase Kinetex 2.6- μ mC₁₈, 100 mm \times 2.1mm (Phenomenex), and the column temperature was maintained at 40 °C. A linear water- acetonitrile (LCMS-grade) gradient was used (both solvents were buffered with 20 mM formic acid) starting from 10% (vol/vol) acetonitrile and increased to 100% in 10min, maintaining this rate for 3 min before returning to the starting conditions in 0.1 min and staying there for 2.4 min before the following run. A flow rate of 0.4 mL·min⁻¹ was used. TOFMS was performed in ESI+ with a data acquisition range of 10 scans per second at m/z 100–1000. The TOFMS was calibrated using Bruker Daltonics high precision calibration algorithm (HPC) by means of the use of the internal standard sodium formate, which was automatically infused before each run. This provided a mass accuracy of better than 1.5 ppm in MS mode. UV/VIS spectra were collected at wavelengths from 200 to 700 nm. Data processing was

performed using DataAnalysis 1.2 and Target Analysis4.0 software (Bruker Daltonics). Tandem MS was performed in with fragmentation energy was varied from 18 to 55 eV.

3. Results and discussion

3.1. PKS44, which has a homologue in *A. carbonarius*, is similar to AlbA

PKS44 (ASPNIDRAFT_191422) is a non-reducing PKS, which has not been linked to production of a compound. In order to investigate the possible polyketide (PK) product of PKS44, a BLAST search was performed on all sequenced Aspergilli in the Aspergillus Genome Database (AspGD). Interestingly, PKS44 had homologues in three other Aspergilli: *A. acidus* (Aspfo1_0143182), *A. brasiliensis* (Aspbr1_0199700), and *A. carbonarius* (Acar5010_056260), which all belong to the group of black Aspergilli. Furthermore, we found that PKS44 was similar to the *A. niger* PKS AlbA (62.3 % similarity/45.5 % identity), and contained the same domain organization SAT-KS-AT-ACP-ACP-TE (Fisch et al., 2009). As AlbA is responsible for production of naphtho- γ -pyrones we speculated that PKS44 might produce a PK similar to the AlbA PK product.

3.2. Heterologous production of three PKSs in *A. nidulans* results in production of a yellow pigment

It has been shown previously that the non-reducing PKS44 contains a thioesterase (TE) domain for product release. As PKSs containing such a TE domain is known to release their polyketide (PK) products without the need for additional enzymes, we expressed the PKS44 gene, its *A. carbonarius* homologue (Acar44), and *albA* heterologously in the model fungus *A. nidulans*, for which a system for heterologous expression has been developed and tested for expression of PKS genes (Hansen et al., 2011, Holm et al., submitted for publication). The PKS genes were expressed from *ISI* under regulation of the strong constitutive promoter *PgpdA*. The resulting strains had visibly altered phenotypes: *ISI*-PKS44 and *ISI*-Acar44 had slightly yellow

colors, which were visible to the eye when looking at the colonies and *ISI-albA* produced vast amounts of yellow pigment, which was secreted into the medium (Figure 2A).

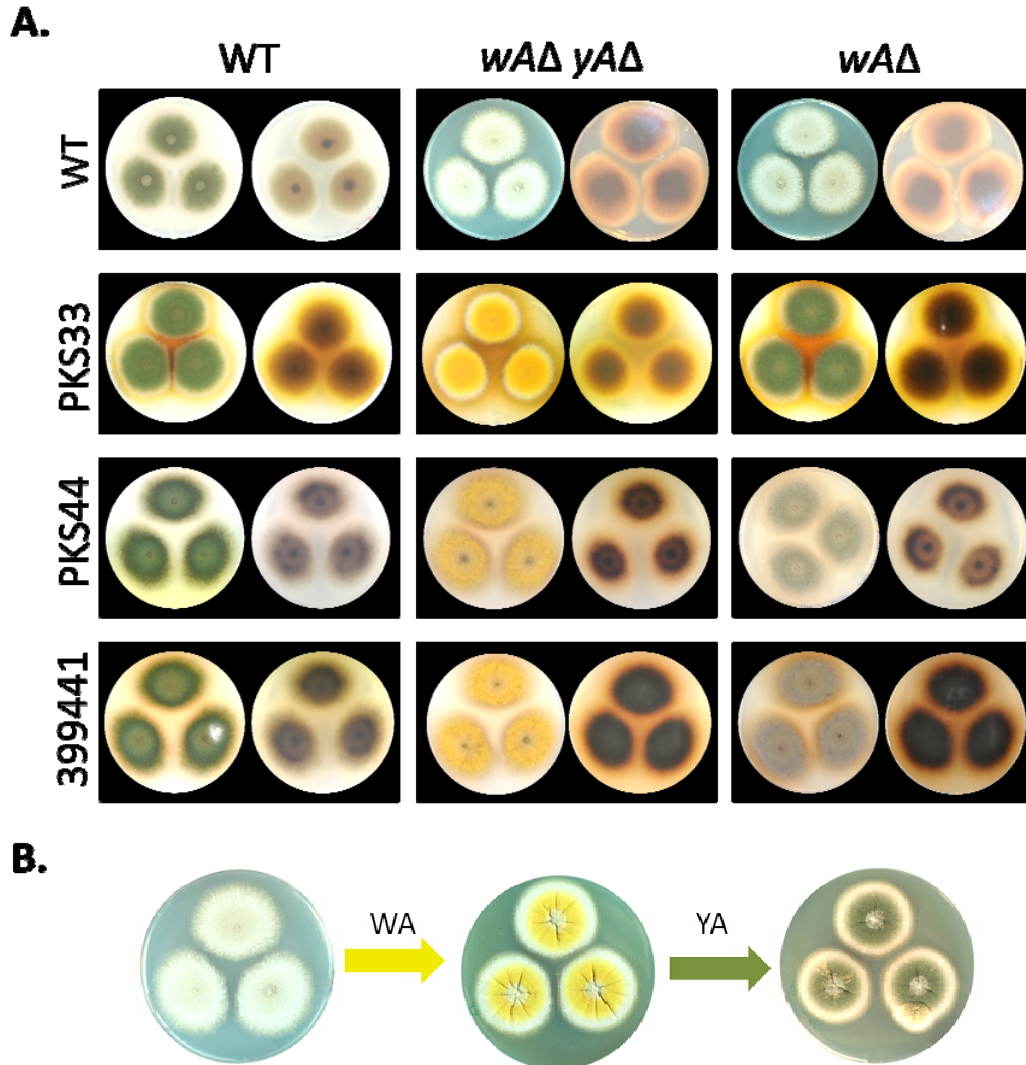


Figure 2. **A.** Morphologies of *A. nidulans* strains expressing *albA*, PKS44, and Acar44. **B.** Conidial pigment pathway in *A. nidulans*. WA synthesizes the yellow pigment YWA1, which is polymerized into green DHN-based melanin by the YA laccase.

The strains were subjected to metabolite profile analysis by UHPLC-DAD-TOFMS and the chromatograms were examined for production of new compounds. Surprisingly, no new compound could be detected in *IS1*-PKS44 or *IS1*-Acar44, but *IS1*-*albA* did produce a new compound (RT = 4.65 min, m/z = 277.0709) relative to the *A. nidulans* reference strain NID1 (Figure 3).

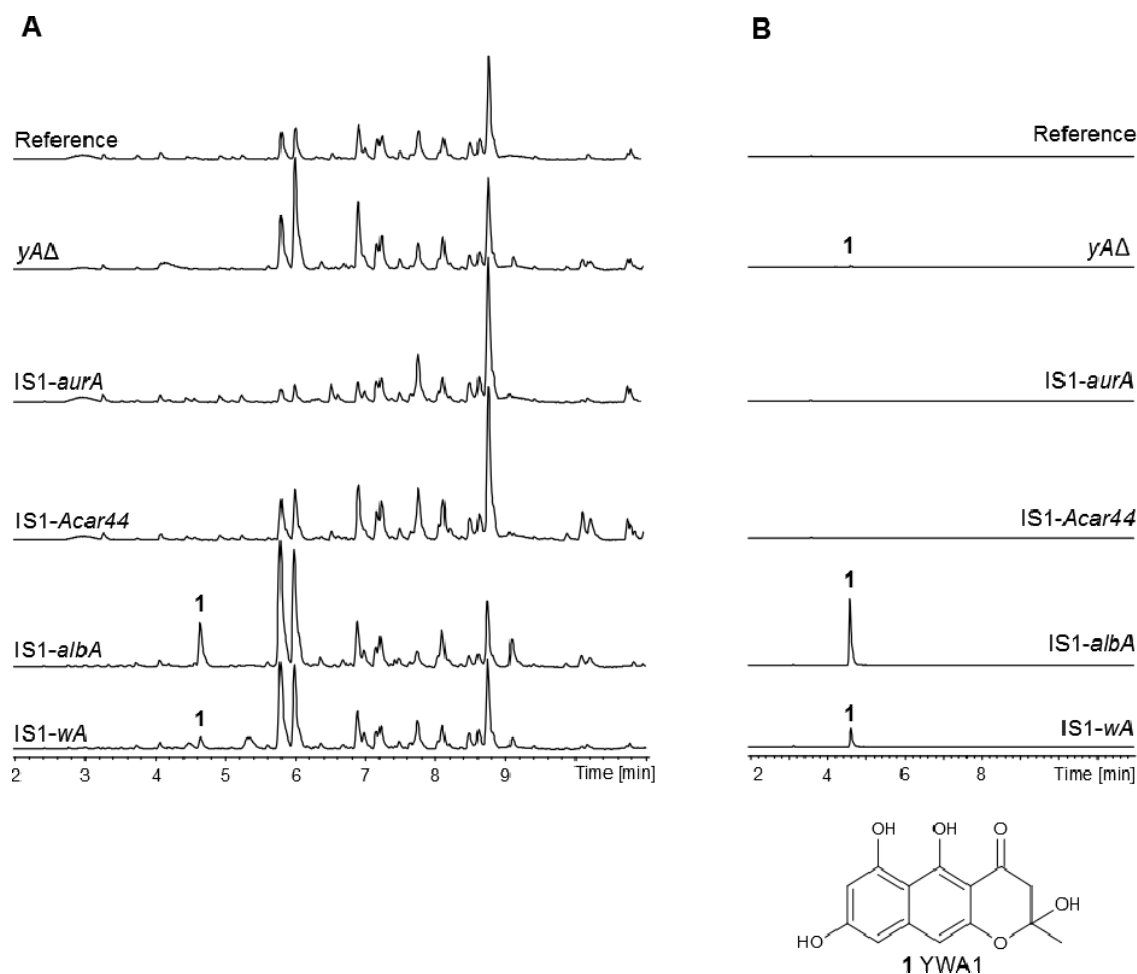


Figure 3. A. BPC of *A. nidulans* strains expressing PKS genes. B. EIC of the pseudomolecular ion of YWA (1) (277.0709 \pm 0.005). Chromatograms are to scale.

Dereplication identified the compound as being YWA1, which has been characterized in *A. nidulans* as the yellow conidial pigment that polymerizes into green DHN-based melanin (Watanabe et al., 1999). The fact that YWA1 can be converted into black DHN-based melanin via chain shortening and subsequent polymerization, combined with the fact that YWA1 is the product of AlbA, led us to conclude that in *A. niger* black melanin is formed from a YWA1 precursor in a similar manner as shown for *A. fumigatus* (Fujii et al., 2004).

Based on the similarity between AlbA and PKS44, we hypothesized that the yellow compound produced by PKS44 in *A. nidulans* might also be a conidial pigment. However, AlbA produced vast amounts of YWA1, which did not seem to be the case for the pale yellow *ISI*-PKS44 strain. Thus, either the PKS44 product is different from YWA1, or the PKS44 had a much lower activity than AlbA. If the compound was bound to insoluble cell material e.g. the cell wall it would certainly explain why it was not detectable by analytical chemistry. To substantiate this hypothesis we analyzed a yellow strain of *A. nidulans* (*yAΔ*) and indeed we detected only trace amounts of YWA1 (Figure 3) although the yellow color was clearly visible to the eye (Figure 2B), showing that the vast majority of the produced YWA1 is impossible to extract.

3.3. Deletion of PKS44 in *A. niger* does not produce a detectable change in the metabolite profile

As heterologous expression of the PKSs did not reveal the PK product(s), we deleted the PKS44 in *A. niger* by replacing the ORF with the *hph* selectable marker. The conidial produced by the PKS44 Δ were still black (Figure 4A), suggesting that PKS44 is not required for production of conidial pigments. Therefore the strain was cultivated on solid YES medium, which is known to support production of many compounds in *A. niger*, and subsequently subjected to metabolite profile analysis by UHPLC-DAD-TOFMS (Figure 4B). The resulting data were minutely examined for compounds that were produced by the reference strain, and not in the PKS44 Δ strain, but no such compound was detected. This result suggests that the gene may be silenced, why a gene deletion would not reveal the identity of the PKS product.

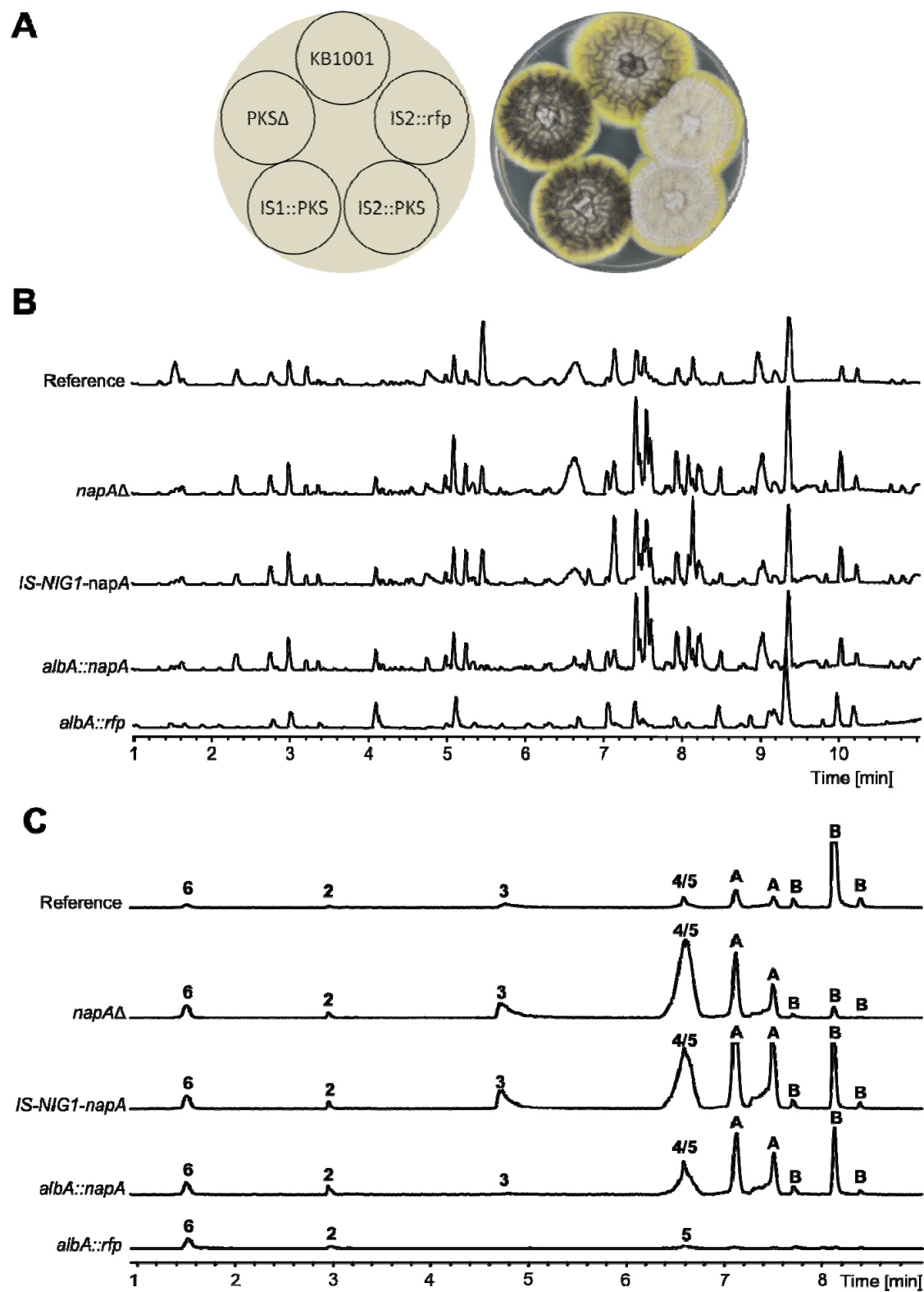


Figure 4. A. Morphologies of *A. niger* strains. B. BPC of *A. niger* strains. C. EIC of *A. niger* strains. A = m/z 589.1704 \pm 0.005, B = m/z 571.1599 \pm 0.005.

3.4. Over-expression of PKS44 in *A. niger albA* results in production of naphtho- γ -pyrones

As the PKS44 seemed to be silent, we over-expressed the PKS44 gene from a specific locus in *A. niger*, *IS-NIG1*, similar to the expression system developed for *A. nidulans* (Hansen et al., 2011). The strain was verified by diagnostic PCR and subjected to metabolite profile analysis by UHPLC-DAD-TOFMS. Still, there was no detectable change in the metabolite profile (Figure 4C). Based on the fact that deletion of PKS44 and over-expression of PKS44 in *A. niger* did not provide any clues about the identity of the PKS product, it was indeed possible that the PKS44 gene and the gene cluster might not be functional at all.

The fact that *albA* produces the yellow YWA1 pigment, which is a precursor to all the naphtho- γ -pyrones in *A. niger*, might result in a masking effect of the compound produced by PKS44, especially if it is produced in small amounts, and/or the activity of the enzyme is low. We therefore over-expressed the PKS44 gene in an *albA* Δ strain to eliminate production of naphtho- γ -pyrones. With only one available marker (*AFpyrG* had been inserted into the *kusA* locus for transient disruption) this was accomplished by integrating the PKS44 gene directly into the *albA* locus, thus looping out *albA*. Correct *albA::PKS44* strains were easily recognizable by the lack of black pigmentation (see Figure 4A), and correct insertion was verified by diagnostic PCR. The strain was subjected to metabolite profile analysis, but there was still no change in the metabolite profile (Figure 4B). This result was extraordinary, as it has been reported twice by different research groups that abolishment of *albA* resulted in the lack of naphtho- γ -pyrones, however, they were still present in this white strain. As the white reference strain (*albA::rfp*) indeed did not produce any naphtho- γ -pyrones (Figure 4C) it only leaves the explanation that PKS44 and AlbA produce the same compound: YWA1.

3.5. The PKS44 product is a substrate to the YA laccase in *A. nidulans*

To test if YWA1 was indeed the product of PKS44, without being able to detect the compound by chemical analysis, we examined the possibility of the PKS44 product being a substrate to the YA laccase, which catalyzes the polymerization of YWA1 into green melanin in *A. nidulans* (see Figure 1B). In order to

examine this possibility we would remove the native YWA1 PKS in *A. nidulans*, WA, and express each of the three putative homologous PKSs in its place. The resulting strains *wA Δ ISI-PKS44*, *wA Δ ISI-Acar44*, and *wA Δ ISI-*albA** were prepared by sexual crossing with a *wA Δ yA Δ* double deletion strain of *A. nidulans*. The strains were examined visually (Figure 1A) and it was clear that the compounds produced by PKS44 and Acar44 were indeed converted into green melanin by the YA laccase, however in a lower concentration than in the reference. As expected, *albA* complemented the *wA* gene completely. These results clearly show that YWA1 is the product of PKS44, why we have named it NapA (Naphtho- γ -pyrone producing), and the corresponding gene *napA*. These results are further substantiated by the fact that Jørgensen and co-workers previously have reported trace amounts of aurasperone B in liquid cultivations of an *albA Δ* strain (Jørgensen et al., 2011b).

3.6. The putative PKS44 gene cluster is identical to the *A. carbonarius* Acar44 gene cluster

Examination of the genes surrounding the PKS44 gene in *A. niger* and Acar44 in *A. carbonarius* showed the presence of two putative transcription factors (TFs), along with several activities, which are typically found in a secondary metabolite cluster: a CYP450, a monooxygenase, two O-methyltransferases, and a dehydrogenase (Figure 5)(Marchler-Bauer et al., 2011). Additionally, the putative cluster in *A. niger* contained two genes with no characterized homologues. Interestingly, the Acar44 gene cluster in *A. carbonarius* was identical to the PKS44 cluster in *A. niger*, although with only one of the genes with no characterized homologues (identities and similarities are shown in Appendix, Table A3). Gene prediction was analyzed using FGeneSH (Softberry), which proposed that the 56024 ORF in *A. carbonarius* was in fact two different ORFs, and a pairwise alignment verified that the two ORFs (termed 56024A and 56024B) were in fact homologues of the two ORFs in *A. niger* (ASPNI_DRAFT191165 and ASPNIDRAFT_191711), respectively. Interestingly, the cluster seems to contain two transcription factors (ASPNIDRAFT_50739 and ASPNIDRAFT_44533), which is also the case for the aurofusarin cluster.

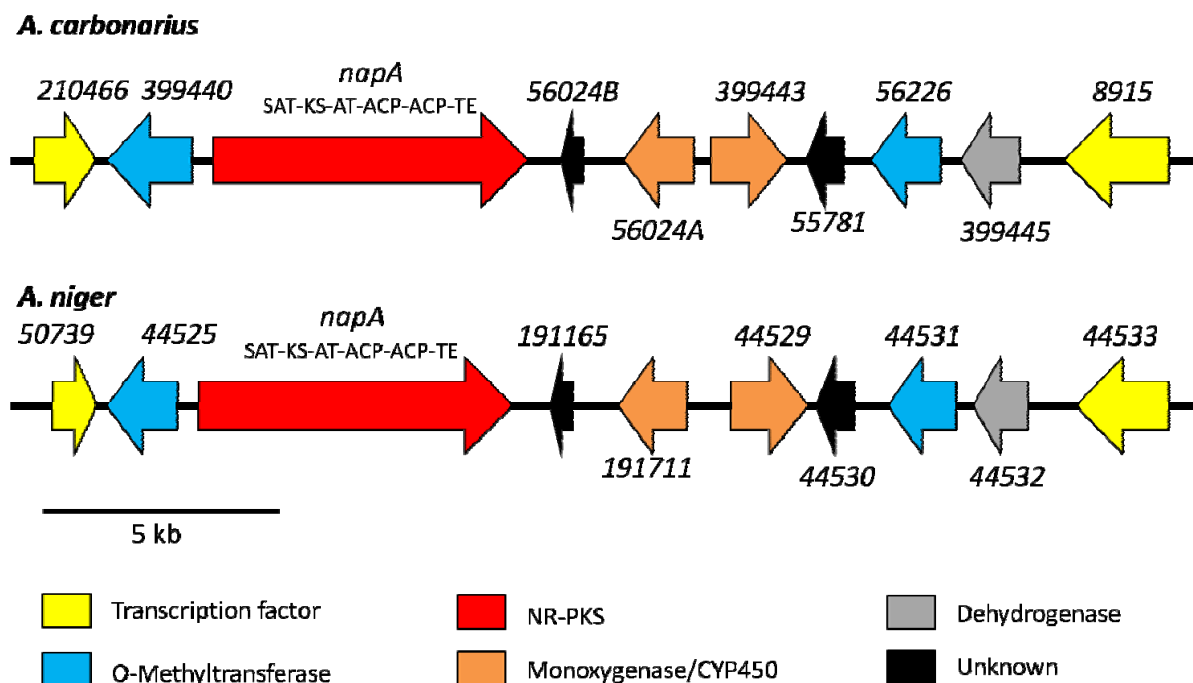


Figure 5. *nap* clusters in *A. carbonarius* and *A. niger*.

The presence of two O-methyltransferases in the gene cluster account for addition of two methyl groups. Both of the O-methyltransferases have homology to AurJ in *Fusarium graminearum* (Frandsen *et al.*, 2006). Furthermore, the enzyme encoded by ASPNIDRAFT_191165 is a homologue of the AurZ dehydrogenase in *F. graminearum* (Frandsen *et al.*, 2011), which catalyzes the dehydration of YWA1 to form nor-rubrofusarin (Figure 6). In *F. graminearum*, nor-rubrofusarin is proposedly oxidized to form 9-hydroxyrubrofusarin, however, this configuration has not been observed in *A. niger* KB1001 (Chiang *et al.*, 2011). In *F. graminearum*, dimerization into aurofusarin is catalyzed by an extracellular complex comprising Gip1 (laccase), AurF, AurO, and AurS. BLAST analysis suggests that these enzymes have no apparent homologues in *A. niger* (ATCC1015), however, a PFAM search (PF00394) revealed that *A. niger* contains no less than 15 putative laccases, none of which are located in the vicinity of *albA* or *napA*. As only O-methylation and dimerization activities are required for formation of the naphtho- γ -pyrones produced in *A. niger*, it is possible that dimerization is catalyzed by all or some of the remaining enzymes encoded by this

not seem to be surrounded by any cluster genes (Chiang et al., 2011), although several genes are involved in chain shortening and modifications of YWA1 into melanin, reviewed in (Langfelder et al., 2003).

However, *napA* is located in a typical gene cluster, containing all the enzymatic activities that are required for modification of YWA1 into naphtho- γ -pyrones. Thus, the *albA* and *napA* genes may originally have been involved in separate processes in different compartments. As the AlbA enzyme has much higher activity than NapA, it may have been energetically favorable to shift to metabolism in a way that allows AlbA to produce YWA1 for feeding into both pathways. The similarity of the two PKS genes suggest that they have arisen via divergent evolution, however, it is not obvious to predict the origin of the two biosyntheses.

Several of the genes encoded by the *napA* gene cluster are similar to genes that have been characterized in the aurofusarin gene cluster in *F. graminearum*. The evolutionary relationship between the two types of clusters is further strengthened by the similarity in products, which both have a YWA1 core that is O-methylated and dimerized via a carbon-carbon bond. Other compounds with similar features include parasperone A from *Aspergillus parasiticus* (Salvo, 1994), ustilaginoidin from *Ustilagoidea virens*, and chaetochromin from *Chaetomium spp.* (Koyama et al., 1988), which likely encode similar biosynthetic gene clusters.

3.8. Conclusions

In this study we have shown that *A. niger* contains two copies of the same gene, *albA* and *napA*, encoding PKSs responsible for production of DHN-melanin and naphtho- γ -pyrones. One of the genes, *napA*, is located in a typical secondary metabolite gene cluster responsible for naphtho- γ -pyrone, but not DHN-melanin, formation. Interestingly, deletion of *albA* abolishes production of naphtho- γ -pyrones, suggesting that *napA* is silenced under laboratory conditions, however, over-expression studies have confirmed that NapA is a functional enzyme and expression of NapA restores naphtho- γ -pyrone production in an *albA* Δ strain. Based on bioinformatics, we have suggested a biosynthetic route towards naphtho- γ -pyrones comprising

dehydration and O-methylation, and we propose that naphtho- γ -pyrones are not formed via laccase-catalyzed dimerization as is the case for aurofusarin in *F. graminearum*, but rather via oxidative phenol coupling as shown for dimerization of bicoumarins in *A. niger*.

Acknowledgments

We kindly acknowledge Kenneth S. Bruno (Pacific Northwest National Laboratory, Richland, WA, USA) for providing the KB1001 strain. The study was supported by grant 09-064967 from the Danish Council for Independent Research, Technology, and Production Sciences.

References

- Bergmann, S, Schümann, J., Scherlach, K, Lange, C., Brakhage, A. A., & Hertweck, C., 2007. Genomics-driven discovery of PKS-NRPS hybrid metabolites from *Aspergillus nidulans*. Nat. Chem. Biol., 3(4), 213-217. doi: 10.1038/nchembio869.
- Bok, J. W., Chiang, Y.-M., Szewczyk, E, Reyes-Dominguez, Y., Davidson, A. D., Sanchez, J. F., et al., 2009. Chromatin-level regulation of biosynthetic gene clusters. Nat. Chem. Biol., 5(7), 462-464.
- Chiang, Y.-M., Szewczyk, E, Davidson, A. D., Keller, N. P., Oakley, B R, & Wang, C. C. C., 2009. A gene cluster containing two fungal polyketide synthases encodes the biosynthetic pathway for a polyketide, asperfuranone, in *Aspergillus nidulans*. JACS, 131(8), 2965-70. doi: 10.1021/ja8088185.
- Chiang, Y.-M., Szewczyk, E, Davidson, A. D., Entwistle, R., Keller, N. P., Wang, C. C. C., et al., 2010. Characterization of the *Aspergillus nidulans* monodictyphenone gene cluster. Appl. Environ. Microb., 76(7), 2067-74. doi: 10.1128/AEM.02187-09.
- Chiang, Y-M., Meyer, K.M., Praseuth, M., Baker, S.E., Bruno, K.S., Wang, C.C.C., 2011. Characterization of a polyketide synthase in *Aspergillus niger* whose product is a precursor for both dihydroxynaphthalene (DHN) melanin and naphtho- γ -pyrone. Fungal Genet. Biol. 48, 430–437.
- Cove, D. J., 1966. The induction and repression of nitrate reductase in the fungus *Aspergillus nidulans*. Biochim. Biophys. Acta. 113, 51-6.
- Fisch, K.M., Gillasp, A.F., Gipson, M., Henrikson, J.C., Hoover, A.R., Jackson, L., Najar, F.Z., Wägele, H., Cichewicz, R.H., 2009. Chemical induction of silent biosynthetic pathway transcription in *Aspergillus niger*. J. Ind. Microbiol. Biotechnol. 36, 1199–1213.
- Frandsen, R.J.N., Nielsen, N.J., Maolanon, N., Sørensen, J.C., Olsson, S., Nielsen, J., Giese, H., 2006. The biosynthetic pathway for aurofusarin in *Fusarium graminearum* reveals a close link between the naphthoquinones and naphthopyrones. Mol. Microbiol. 61,1069–1080.

- Frandsen, R.J.N., Schütt, C., Lund, B.W., Staerk, D., Nielsen, J., Olsson, S., Giese, H., 2011 Two novel classes of enzymes are required for the biosynthesis of aurofusarin in *Fusarium graminearum*. J. Biol. Chem. 286,10419-10428.
- Frisvad, J.C., Samson, R.A., 2004. Polyphasic taxonomy of *Penicillium* subgenus *Penicillium*. A guide to identification of food and air-borne terverticillate *Penicillia* and their mycotoxins. Stud. Mycol. 49, 1–173.
- Fujii, I., Yasuoka, Y., Tsai, H.F., Chang, Y.C., Kwon-Chung, K.J., Ebizuka, Y., 2004. Hydrolytic polyketide shortening by *ayglp*, a novel enzyme involved in fungal melanin biosynthesis. J. Biol. Chem., 279:44613-44620.
- Geu-Flores, F., Nour-Eldin, H.H., Nielsen, M.T., Halkier, B.A., 2007. USER fusion: a rapid and efficient method for simultaneous fusion and cloning of multiple PCR products. Nucleic Acids Res. 35, 7.
- Gil Girol, C., Fisch, K.M., Heinekamp, T., Günther, S., Hüttel, W., Piel, J., Brakhage, A.A., Müller, M. 2012. Regio- and Stereoselective Oxidative Phenol Coupling in *Aspergillus niger*. Angew. Chem. Int. Ed. 51, 9788 –979.
- Hansen, B.G., Salomonsen, B., Nielsen, M.T., Nielsen, J.B., Hansen, N.B., Nielsen, K.F., Regueira, T.B., Nielsen, J., Patil, K.R., Mortensen, U.H., 2011. Versatile Enzyme Expression and Characterization System for *Aspergillus nidulans*, with the *Penicillium brevicompactum* Polyketide Synthase Gene from the Mycophenolic Acid Gene Cluster as a Test Case. Appl. Environ. Microb. 77, 3044–3051.
- Holm, D.K., Petersen, L.M., Klitgaard, A., Knudsen, P.B., Jarczynska, Z.D., Nielsen, K.F., Gotfredsen, C.H., Larsen, T.O., Mortensen, U.H. Molecular and chemical characterization of the biosynthesis of the 6-MSA derived meroterpenoid yanuthone D in *Aspergillus niger*. JACS, submitted for publication.
- Johnstone, I.L., Hughes, S.G. and Clutterbuck, A.J., 1985. Cloning an *Aspergillus nidulans* developmental gene by transformation. EMBO J. 4, 1307–1311.

Koyama, K., Ominato, K., Natori, S., Tashiro, T., Tsuruo, T. 1988. Cytotoxicity and Antitumor Activities of Fungal Bis (naphtho- γ -pyrone) Derivatives. J. PHARMACOBIO-DYNAM. 11, 9, 630-635.

Langfelder, K., Streibel, M., Jahn, B., Haase, G., Brakhage, A.A., 2003. Biosynthesis of fungal melanins and their importance for human pathogenic fungi. Fungal Genet. Biol. 38, 143–158.

Jørgensen, T.R., Park, J., arentshorst, M., van Welzen, A.M., Lamers, G., van Kuyk, P.A., Damveld, R.A., van den Hondel, C.A.M., Nielsen, K.F., Frisvad, J.C., Ram, A.F., 2011a. The molecular and genetic basis of conidial pigmentation in *Aspergillus niger*. Fungal Genet. Biol. 48, 544–553

Jørgensen, T.R., Nielsen, K.F., Arentshorst, M., Park, J., van den Hondel, C.A., Frisvad, J.C., Ram, A.F., 2011b. Submerged Conidiation and Product Formation by *Aspergillus niger* at Low Specific Growth Rates Are Affected in Aerial Developmental Mutants. Appl. Environ. Microb. 5270–5277.

doi:10.1128/AEM.00118-11.

Marchler-Bauer, A., Lu, S., Anderson, J.B., Chitsaz, F., Derbyshire, M.K., DeWeese-Scott, C., Fong, J.H., Geer, L.Y., Geer, R.C., Gonzales, N.R., Gwadz, M., Hurwitz, D.I., Jackson, J.D., Ke, Z., Lanczycki, C.J., Lu, F., Marchler, G.H., Mullokandov, M., Omelchenko, M.V., Robertson, C.L., Song, J.S., Thanki, N., Yamashita, R.A., Zhang, D., Zhang, N., Zheng, C., Bryant, S.H., 2011. . CDD: a Conserved Domain Database for the functional annotation of proteins. Nucleic Acids Res. 39, 225-9.

McCluskey, K., Wiest, A., Plamann, M., J., 2010. The Fungal Genetics Stock Center: a repository for 50 years of fungal genetics research. Biosci., 35: 119–126.

Nielsen, K.F., Mogensen, J.M., Johansen, M., Larsen, T.O., Frisvad, J.C., 2009. Review of secondary metabolites and mycotoxins from the *Aspergillus niger* group. Anal. Bioanal. Chem. 395, 1225-1242.

Nielsen, M.L., Albertsen, L., Lettier, G., Nielsen, J.B., Mortensen, U.H., 2006. Efficient PCR-based gene targeting with a recyclable marker for *Aspergillus nidulans*. Fungal Genet. Biol. 43, 54-64.

Nørholm, M.H.H., 2010. A mutant Pfu DNA polymerase designed for advanced uracil excision DNA engineering. *BMC Biotechnol.* 10, 21.

Olsen, L.R., Hansen, N.B., Bonde, M.T., Genée, H.J., Holm, D.K., Carlsen, S., Hansen, B.G., Patil, K.R., Mortensen, U., Wernersson R., 2011. PHUSER (Primer Help for USER): A novel tool for USER fusion primer design. *Nucleic Acids Res.* 1–7. doi:10.1093/nar/gkr394.

Salvo, J.J. 1994. Structural Elucidation of a Putative Conidial Pigment Intermediate in *Aspergillus parasiticus*. *Tetrahedron Letters*, 34, 3, 419-422.

Schroeckh, V., Scherlach, K., Nützmann, H.-W., Shelest, E., Schmidt-Heck, W., Schuemann, J., et al., 2009. Intimate bacterial-fungal interaction triggers biosynthesis of archetypal polyketides in *Aspergillus nidulans*. *PNAS*, 106(34), 14558-63. doi: 10.1073/pnas.0901870106.

Smedsgaard, J., 1997. Micro-scale extraction procedure for standardized screening of fungal metabolite production in cultures. *J. Chromatogr. A* 760, 264-270.

Szewczyk, E., Chiang, Y.-M., Oakley, C. E., Davidson, A. D., Wang, C. C. C., & Oakley, B. R., 2008. Identification and characterization of the asperthecin gene cluster of *Aspergillus nidulans*. *Appl. Environ. Microb.*, 74(24), 7607-7612. doi: 10.1128/AEM.01743-08.

Watanabe, A., Fujii, I., Sankawa, U., Mayorga, M.E., Timberlake, W.E., Ebizuka, Y., 1999. Re-identification of *Aspergillus nidulans* *wA* Gene to Code for a Polyketide Synthase of Naphthopyrone. *Tetrahedron Lett.* 40, 91-94.

Chapter 10

Molecular and chemical characterization of the biosynthesis of the 6-MSA derived meroterpenoid yanuthone D in *Aspergillus niger*

This chapter contains the manuscript "Molecular and chemical characterization of the biosynthesis of the 6-MSA derived meroterpenoid yanuthone D in *Aspergillus niger*", submitted to *JACS*, of which I am joint first author.

In this interdisciplinary study I have dissected the biosynthetic pathway towards yanuthone D in *A. niger* by combining several disciplines within molecular biology, preparative chemistry and analytical chemistry including gene deletions, gene expressions using the newly developed AMA1-based expression platform (described in Section 3.1.4), structure elucidation by NMR spectroscopy, in-house production of fully labeled 6-methylsalicylic acid using one of the newly developed integration sites (described in Section 3.1.2), identification of compounds by MS/MS, and feeding studies with both labeled and unlabeled compounds.

My contribution to this study was conception of the idea based on results from Chapter 8, preparation of strains in collaboration with Z.D.J., planning and execution of feeding studies and analysis thereof, and writing of the manuscript in collaboration with L.M.P. and the entire team. Supporting information is included in Appendix E.

Molecular and chemical characterization of the biosynthesis of the 6-MSA derived meroterpenoid yanuthone D in *Aspergillus niger*

Dorte K. Holm^{†,1}, Lene M. Petersen^{†,2}, Andreas Klitgaard³, Peter B. Knudsen⁵, Zofia D. Jarczyska¹, Kristian F. Nielsen³, Charlotte H. Gotfredsen⁴, Thomas O. Larsen^{2,*}, Uffe H. Mortensen^{1,*}

Affiliation:

¹Eukaryotic Molecular Cell Biology group, Center for Microbial Biotechnology, Department of Systems Biology, Soltofts Plads, Building 223, Technical University of Denmark, 2800 Kgs. Lyngby, Denmark.

²Chemodiversity group, Center for Microbial Biotechnology, Department of Systems Biology, Soltofts Plads, Building 221, Technical University of Denmark, 2800 Kgs. Lyngby, Denmark.

³Metabolic Signaling and Regulation group, Center for Microbial Biotechnology, Department of Systems Biology, Soltofts Plads, Building 221, Technical University of Denmark, 2800 Kgs. Lyngby, Denmark.

⁴Department of Chemistry, Kemitorvet, Building 201, Technical University of Denmark, 2800 Kgs. Lyngby, Denmark.

⁵Fungal Physiology and Biotechnology group, Center for Microbial Biotechnology, Department of Systems Biology, Soltofts Plads, Building 223, Technical University of Denmark, 2800 Kgs. Lyngby, Denmark.

†These authors have contributed equally to this work.

***Corresponding authors:**

Uffe H. Mortensen, Technical University of Denmark, Department of Systems Biology
Center for Microbial Biotechnology, Eukaryotic Molecular Cell Biology, Building 223
2800 Kgs. Lyngby, Denmark

Email: um@bio.dtu.dk

Phone: +45 45252701

Fax: +45 45884148

Thomas O. Larsen, Technical University of Denmark, Department of Systems Biology

Center for Microbial Biotechnology, Natural Product Chemistry, Building 221

2800 Kgs. Lyngby, Denmark

Email: tol@bio.dtu.dk

Phone: +45 45252632

Fax: +45 45884922

ABSTRACT: Meroterpenoids are produced by several genera of filamentous fungi. In this study we show that the antibiotic yanuthone D from *Aspergillus niger* is a meroterpenoid derived from the polyketide 6-methylsalicylic acid (6-MSA). Formation of yanuthone D depends on a cluster composed of 10 genes including *yanA* encoding the polyketide synthase (PKS) responsible for production of 6-MSA. All genes in this cluster were individually deleted and analyses of the resulting strains uncovered three intermediates towards yanuthone D. These findings have allowed us to deduce the genetic and biosynthetic pathway from 6-MSA to yanuthone D including linking one of the genes, *yanI*, to a novel enzymatic activity. In addition, several branching points in the pathway were discovered, revealing five novel yanuthones (F, G, H, I, and J) that are all likely the result of detoxification processes. Identification of another novel compound (yanuthone X₁) revealed a structure similar to the other yanuthones, however, despite the striking similarity to the yanuthones, both gene deletions and ¹³C₈-6-MSA feeding experiments showed that yanuthone X₁ does not originate from 6-MSA. Rather it is synthesized via a different precursor, which is nevertheless modified in a process that is dependent on several genes in the *yan* cluster, thus defining a novel class of yanuthones.

INTRODUCTION

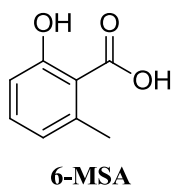
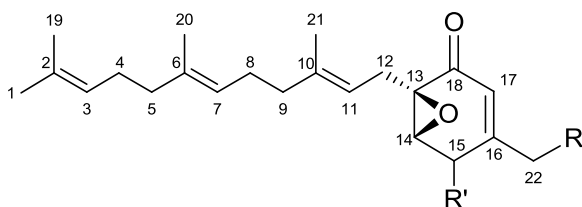
Fungal polyketides (PKs) comprise a large and complex group of metabolites with a wide range of bioactivities. Hence, the group includes compounds that are used by fungi as pigments for UV-light protection, in intra- and inter species signalling and in chemical warfare against competitors¹. Many PKs are mycotoxins that are harmful to human health e.g. patulin and the highly carcinogenic aflatoxins². On the other hand, several PKs have a great medical potential e.g. cholesterol-lowering statins³, the antimicrobial and immunosuppressive mycophenolic acid⁴, the ACAT inhibiting pyripyropenes⁵, and the farnesyltransferase inhibiting andrastins⁶. Although more than 6000 different PKs have been isolated and characterized (Antibase 2012), these compounds are likely only the tip of the iceberg. For example, for each fungus analyzed, only a small part of its full repertoire of PKs genes appears to be produced under laboratory conditions^{7,8}. In agreement with this view, genome sequencing of several fungal species have uncovered far more genes for PKs production than can be accounted for by the number of compounds that they are actually known to produce. Hence, the chemical space of PKs is far from fully known and many new drugs and mycotoxins await discovery.

The fungal genome sequencing projects have demonstrated that genes necessary for production of individual PKs often cluster around the gene encoding the polyketide synthase (PKS), which delivers the first intermediate in a given PK pathway. Although, this is helpful for pathway elucidation, compounds produced by orphan gene clusters⁹ can still not be easily predicted by bioinformatic tools, reviewed in^{10,11}. This is because most fungal PKs are produced by type I iterative PKSs whose products are notoriously difficult to predict. Moreover, the specificities and the order of actions of the tailoring enzymes that modify the PK released from the PKS further complicate prediction of the end products. To elucidate the biochemical pathway of an orphan gene cluster it is therefore necessary to create gene cluster mutations and/or genetically reconstitute the pathway in a heterologous host. Subsequent analytical and structural chemistry analyses of the compounds that are present in the reference strain, but not in the mutant strains; and of compounds that accumulate in the mutant

strains, but are absent or present in minute amounts in the reference strain, may deliver insights that can be used for pathway elucidation.

Aspergillus niger is an industrially important filamentous fungus, which has obtained GRAS status for use in several industrial processes and is used for production of organic acids and enzymes. Importantly, when the full genome sequence of *A. niger* was examined, a gene cluster resembling the fumonisins gene cluster from *Gibberella moniliformis* was surprisingly identified, suggesting that this well characterized fungus had the genetic potential to produce the carcinogenic fumonisins¹². This possibility was later confirmed by genetic and chemical analyses^{7,13}. The fact that the *A. niger* genome contains several orphan gene clusters for production of secondary metabolites¹⁴ raises the question, whether it can produce other bioactive PKs that could be harmful, or perhaps beneficial, to human health. To this end, it is interesting to note that among the 33 predicted PKS and PKS-like genes in *A. niger*, one encodes a putative PKS, which is phylogenetically close to fungal 6-methylsalicylic acid (6-MSA) synthases¹⁴. Importantly, the model PK 6-MSA¹⁵ (Figure 1A) is known to be the precursor to e.g. the mycotoxin patulin¹⁶ produced by many *Aspergillus* and *Penicillium* species substantiating the possibility that this gene could be the source of yet another unknown bioactive PK in *A. niger*. We therefore investigated whether *A. niger* has the potential to produce 6-MSA or 6-MSA derived compounds.

Yanuthones constitute a group of compounds that comprise a six-membered methylated ring (the C₇ core scaffold) with two side chains; one sesquiterpene and varying side chain (-R) (Figure 1B). In this study we demonstrate that in *A. niger*, 6-MSA is the precursor for formation of yanuthone D, which is an antibiotic against *Candida albicans*, methicillin resistant *Staphylococcus aureus* (MRSA) and vancomycin resistant *Enterococcus*¹⁷. We also show that yanuthone D is in fact a complex meroterpenoid synthesized by a pathway where 6-MSA is decarboxylated, heavily oxidized and fused to a sesquiterpene and a mevalon moiety. This is surprising as yanuthones have so far been thought to originate from the shikimate pathway¹⁷.

A**B**

Compound	R	R'
Yanuthone A	OAc	◀OH, ▯H
Yanuthone B	OAc	=O
Yanuthone C	OH	◀OAc, ▯H
Yanuthone D (1)		=O
Yanuthone E (2)		◀OH, ▯H
7-deacetoxyyanuthone A (3)	H	◀OH, ▯H
22-deacetylyanuthone A (6)	OH	◀OH, ▯H

Figure 1. Chemical structures of (A) 6-methylsalicylic acid (6-MSA) and (B) previously described yanuthones: yanuthones A-E, 7-deacetoxyyanuthone A and 22-deacetylyanuthone A^{17,18}.

RESULTS

A. niger PKS48 encodes a 6-MSA synthase. To investigate the possibility that the *A. niger* gene PKS48/ASPNIDRAFT_44965 encodes a 6-MSA synthase, we transferred the gene to *A. nidulans*, which has not

been shown to produce 6-MSA, and which does not contain a close homologue to known 6-MSA PKSs. To ensure a high expression level on a defined medium, the PKS48 gene was integrated into a well characterized integration site, *IS1*¹⁹, under control of the strong constitutive promoter *PgpdA*. As expected, the metabolite profile obtained with an *A. nidulans* reference strain did not show any indications of 6-MSA when analyzed by HPLC-DAD-TOFMS (Figure 2A). In contrast, the metabolite profile of the strain expressing PKS48 showed the presence of a prominent new peak, which had the same retention time as an authentic 6-MSA standard and displayed the same adducts and monoisotopic mass for the pseudomolecular ion. We therefore conclude that PKS48 encodes a 6-MSA synthase.

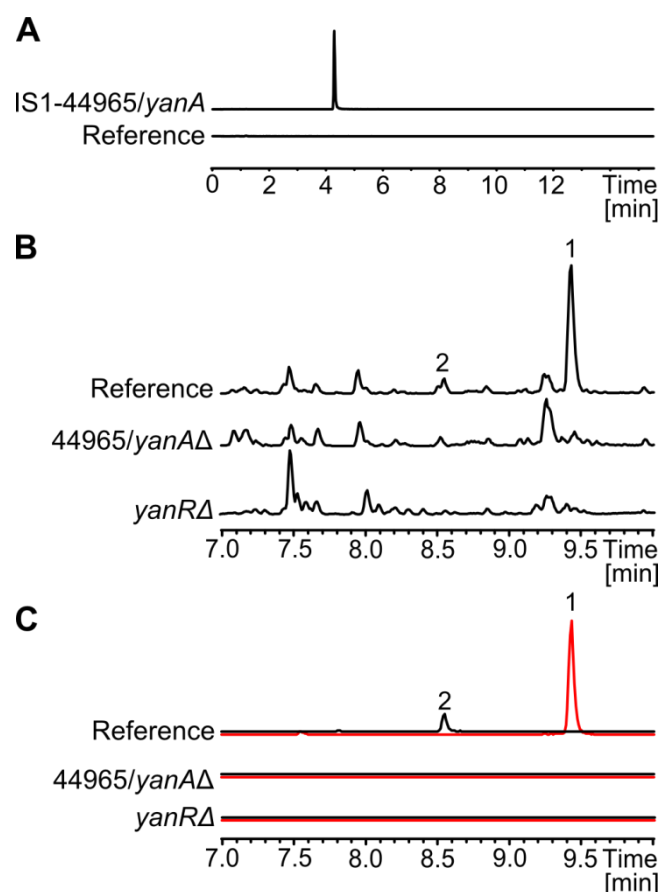


Figure 2. (A) Extracted ion chromatogram (EIC) showing production of 6-MSA (m/z 153.0546 \pm 0.005) in *A. nidulans*. (B) Base peak chromatograms (BPC) m/z 100-1000 of the reference, *yanA* Δ , and *yanR* Δ strain, and (C)

EICs of yanuthone D (**1**) 503.2640 ± 0.005 (red) and yanuthone E (**2**) 505.2791 ± 0.005 (black) for the reference, *yanA* Δ , and *yanR* Δ . All chromatograms are to scale.

Production of yanuthones D and E is eliminated by deletion of PKS48. The fact that 6-MSA has not previously been reported from *A. niger* prompted us to investigate whether this compound could be a precursor to a known secondary metabolite produced by this fungus. We therefore cultivated a reference and a PKS48 Δ *A. niger* strain on four different solid media (MM, CYA, YES and MEA) that are known to trigger the production of a wide range of metabolites²⁰. The resulting UHPLC-DAD-TOFMS metabolite profiles were almost identical (see Figure S1) showing that the PKS48 Δ mutation did not induce a global response on the secondary metabolism. However, on YES and MM media, we identified two compounds that were produced by the reference strain, but not by the PKS48 Δ strain (Figure 2B and C, Table S1). UHPLC separation with UV/Vis and high-resolution MS detection as well as MS/MS suggested that the two compounds were yanuthones D and E. This was confirmed by isolation of the compounds, NMR spectroscopy and circular dichroism (CD) (Supplemental Tables S2, S3 and S4). Hence, production of yanuthones D and E appears to be based on the use of 6-MSA as a key precursor. In this scenario, one carbon must be eliminated from C₈-based 6-MSA to form the C₇ core scaffold of yanuthones D and E.

Yanuthones constitute a complex group of compounds that appear to origin from different precursors. In addition to yanuthones D and E, *A. niger* has previously been reported to produce yanuthones A, B, and C, 1-hydroxyyanuthone A, 1-hydroxyyanuthone C, and 22-deacetylyanuthone A¹⁷ and 7-deacetoxyyanuthone A has been reported from the genus *Penicillium*¹⁸ (Figure 1B). We thus examined the extracted ion chromatograms from the UHPLC-DAD-TOFMS profiles obtained by the reference strain for the presence of these metabolites. In extracts obtained after cultivation on MM, YES, and CYA media, this analysis identified trace amounts of a compound (yanuthone X₁) with a mass and elemental composition corresponding to the yanuthone isomers A and C. The nature of this compound was further investigated by MS/MS and its fragmentation pattern was

similar to the pattern of other yanuthones, showing characteristics such as loss of a sesquiterpene chain. Moreover, the UV/VIS spectrum of the compound was similar to spectra obtained for yanuthones D and E substantiating that this compound was a yanuthone. Surprisingly, when the UHPLC-DAD-TOFMS metabolite profiles obtained with the PKS48 Δ strain were examined for the presence of this yanuthone, it was still present. This observation strongly suggested that some yanuthones, are produced independently of PKS48.

Fully labeled $^{13}\text{C}_8$ -6-MSA is incorporated into yanuthones D and E *in vivo*. The fact that some yanuthones could be produced independently of PKS48 combined with the fact that yanuthones have been proposed to origin from the shikimate pathway, raised the possibility that the absence of yanuthones D and E in the PKS48 deletion strain potentially could be the result of an indirect effect. To investigate this possibility, we fed fully labeled $^{13}\text{C}_8$ -6-MSA to a reference strain and the PKS48 Δ strain at different time points during growth (24 h, 48 h and 72 h), see Experimental Section. The addition of $^{13}\text{C}_8$ -6-MSA did not seem to adversely affect the growth rate, and the morphologies of the colonies of the two strains were identical (Figure S2). This indicates that the amounts of $^{13}\text{C}_8$ -6-MSA added (2-10 $\mu\text{g/ml}$) did not significantly influence strain fitness. Metabolites were then extracted from the plates and analyzed by UHPLC-DAD-TOFMS. For both strains, $^{13}\text{C}_8$ -6-MSA was incorporated into yanuthones D and E resulting in a mass shift of 7.023 Da. This is in agreement with the scenario described above where one carbon atom must be eliminated from 6-MSA in the biosynthetic processing towards yanuthones D and E. Moreover, the MS-based metabolite profiles also showed that $^{13}\text{C}_8$ -6-MSA was exclusively incorporated into compounds related to yanuthones. These compounds are only present in tiny amounts, and are likely intermediates or analogues of yanuthone D or E as they share the same UV chromophore, and as their masses corresponded to water loss(es) or gain from yanuthone D or E. Based on these results, we named the 6-MSA synthase (encoded by PKS48/ASPNIIDRAFT_44965) YanA (yanuthone) and the corresponding gene *yanA*. On the other hand, no labeled yanuthone X₁ was observed in the reference strain as well as in the PKS48 deletion strain after addition of $^{13}\text{C}_8$ -6-MSA (mass spectra are shown in Figure S3),

confirming our finding that yanuthone X₁ is formed in the absence of PKS48. Hence, we conclude that 6-MSA is not the precursor of yanuthone X₁.

The *yan* gene cluster comprises ten genes. To determine whether *yanA* defines a gene cluster for a biosynthetic pathway towards yanuthones D and E, ten genes up- and downstream of *yanA* were annotated using FGeneSH (Softberry) and AUGUSTUS software²¹. Subsequently, these twenty putative genes were examined using the NCBI Conserved Domain Database (CDD)²² for ORFs encoding activities that are typically employed for the modification of PKs. Based on these analyses, additional eight genes could potentially belong to the *yanA* cluster, including genes encoding a transcription factor (TF), a prenyl transferase, an O-acyl transferase, a decarboxylase, two oxidases, two cytochrome P450s (CYP450s), and a dehydrogenase (Table 1 and Figure 3). Together with *yanA* and 192604 (a gene with no known homologs), these eight genes form a cluster of ten genes that are not interrupted by any of the remaining eleven genes included in the analysis. The fact that one of the ten genes in this cluster putatively encodes a TF (44961) raised the possibility that expression of the genes involved in yanuthones D and E production is controlled by this TF. In agreement with this view, deletion of 44961 resulted in a strain that did not produce these two yanuthones (Figure 2B and C). To further delineate the *yanA* gene cluster, we determined the expression levels of the ten cluster genes as well as of four flanking genes by RT-qPCR in a 44961Δ strain and a reference strain. When the two datasets were compared, we found, as expected, that expression from 44961 is eliminated in the 44961Δ strain where the entire gene is deleted (see Figure S4). More importantly, the analysis demonstrated that expression from the other nine genes in the cluster was significantly down-regulated in the 44961Δ strain as compared to the reference strain (*p*-value < 0.05). Specifically, the expression was reduced more than 10-fold for seven of the genes, including *yanA*. Expression of the remaining two genes: 54844 and 44964 were expressed at a level corresponding to 20 % and 11 %, respectively, of the level obtained with the reference strain. In contrast, expression levels from the four flanking genes were not significantly different from the reference (Figure S4). Next, we individually deleted the remaining eight genes in the proposed *yan* gene cluster, which encode putative activities for PK modification.

None of the resulting strains, including 192604 Δ , produced yanuthone D indicating that all genes belong to the *yan* cluster (Table S1). As a control, the four additional genes flanking this cluster were also individually deleted, but all these four strains produced yanuthone D. Based on these analyses and the results from the RT-qPCR, we propose that the *yan* gene cluster is composed by 10 genes *yanA-I*, and *yanR*, where *yanR* encodes a TF that regulates the gene cluster (Figure 3, Table 1). Finally, all ten genes were simultaneously deleted in one strain. When $^{13}\text{C}_8$ -6-MSA was fed to this strain, no labeled metabolites were detected showing that all 6-MSA derived yanuthones depend on this gene cluster (see above).

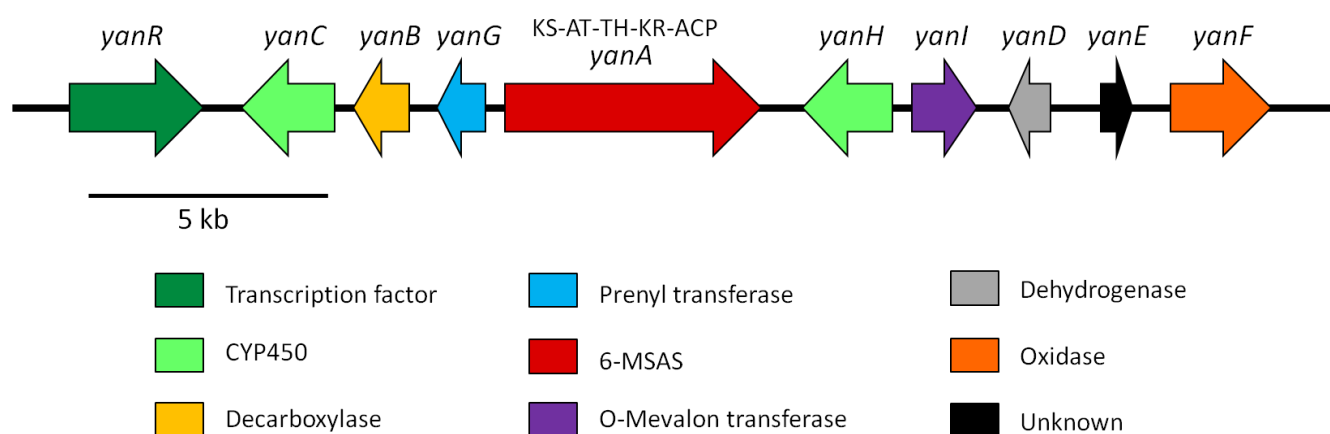


Figure 3. The proposed *yan* cluster. The *yanA* 6-MSA synthase encoding gene is flanked by 9 cluster genes (*yanB*-*yanI*, *yanR*) whose products contain all necessary activities for conversion of 6-MSA into yanuthone D.

Table 1. Overview of genes and proposed activities of the *yan* cluster.

Locus (ASPENIDRAFT_)	Gene name	Predicted functional domains (CDD)	Proposed activity
44959	-	No conservation	-
44960	-	Glyoxalase	-
44961	<i>yanR</i>	Fungal transcription factor	Transcription factor
54844	<i>yanC</i>	CYP450	<i>m</i> -Cresol hydroxylase
44963	<i>yanB</i>	Amidohydrolase, decarboxylase	6-MSA decarboxylase
44964	<i>yanG</i>	UbiA-like prenyltransferase	Prenyltransferase
44965	<i>yanA</i>	Polyketide synthase	6-MSA synthase

193092	<i>yanH</i>	CYP450	Cytochrome P450
44967	<i>yanI</i>	Membrane bound O-acyl transferase	O-Mevalon transferase
127904	<i>yanD</i>	Short-chain dehydrogenase	Dehydrogenase
192604	<i>yanE</i>	No conservation	-
44970	<i>yanF</i>	FAD/FMN-containing dehydrogenases	Oxidase
44971	-	No conservation	-
44972	-	No conservation	-

YanF converts yanuthone E into yanuthone D. As the first step towards elucidating the order of reaction steps in the pathway towards yanuthones D and E, we asked whether yanuthones D and E are two different end products, or whether one is an intermediate in the pathway towards production of the other. To this end, we note that individual deletion of genes in the *yan* gene cluster generally resulted in loss of production of both yanuthones D and E on YES medium. The only exception is the *yanF* Δ strain, which produced substantial amounts of yanuthone E (2), but no yanuthone D (1), see Figure 4. These finding suggest that YanF converts yanuthone E into yanuthone D, which is the true end product of the pathway. Interestingly, the *yanF* Δ strain produced a new and unknown compound, which was not detected in the reference strain. Elucidation of its structure revealed a yanuthone E analog with a hydroxylation at C-2 at the expense of the first double bond (between C-2 and C-3) in the sesquiterpene moiety (Table S5). This compound was named yanuthone J (9).

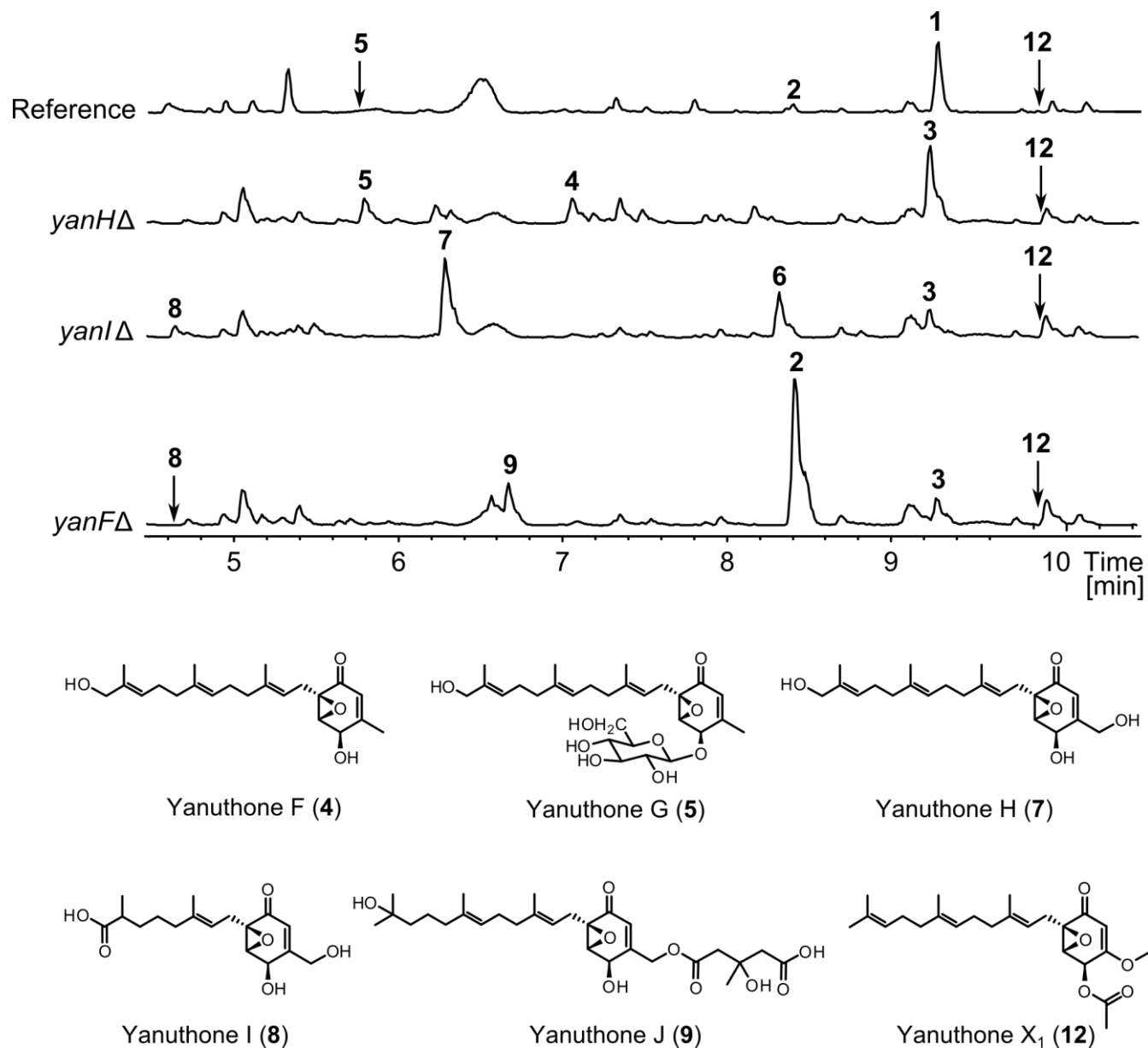


Figure 4. BPC m/z 100-1000 of reference strain (KB1001), *yanHΔ*, *yanIΔ*, and *yanFΔ*. All NMR-elucidated compounds are shown for comparison of intensity and relative retention times. Below are structures of the novel yanuthones identified in this study. Structures of yanuthone D (1), yanuthone E (2), 7-deacetoxyyanuthone A (3), and 22-deacetyyanuthone A (6) are shown in Figure 1.

***m*-cresol and toluquinol are intermediates of the yanuthone D biosynthesis.** Deletion of *yanB*, *yanC*, *yanD*, *yanE*, and *yanG*, did not produce any detectable intermediates and the phenotype of these mutations do therefore not link any of the genes to specific reaction steps in the pathway towards formation of yanuthone D. However, one of the five putative enzymes, YanC, has a defined homologue, PatI, in the *Aspergillus clavatus* patulin biosynthesis pathway²³ where it catalyzes the oxidation of *m*-cresol into toluquinol, suggesting that toluquinol and *m*-cresol are also likely intermediates in the yanuthone biosynthesis. To test this hypothesis we fed *m*-cresol and toluquinol to the *yanAΔ* strain. Analysis of the metabolite profiles of the two strains indeed showed that addition of *m*-cresol or toluquinol restored production of yanuthones D and E in the *yanAΔ* strain (Figure 5).

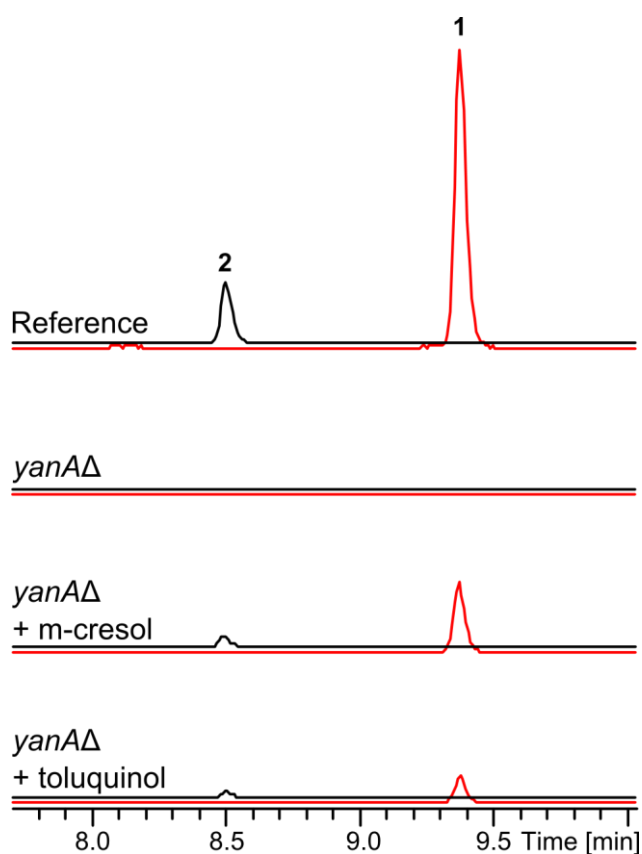


Figure 5. Feeding with unlabeled *m*-cresol and toluquinol. EICs of yanuthone D (**1**) 503.2640 ± 0.005 (red) and yanuthone E (**2**) 505.2791 ± 0.005 (black) for the wildtype and the *yanAΔ* strain with and without feeding. Chromatograms are to scale.

In an attempt to further elucidate the role of the five enzymes, the corresponding genes were inserted into plasmid pDHX2 (Figure S5) and individually expressed in the *A. nidulans* strain harboring the *yanA* gene. No new compounds were produced in these *ISI-yanA* strains expressing *yanC*, *yanD*, *yanE*, and *yanG*, despite that 6-MSA was produced in high amounts (Figure S6). Similarly, in the strain expressing *yanB*, no new product was observed, but in this case 6-MSA was absent indicating that 6-MSA is a substrate for YanB.

Deletion of *yanI* and *yanH* reveals key intermediates in the biosynthesis of yanuthone D. In contrast to the *yanBΔ-EΔ*, and *yanGΔ* strains, new products were observed in the *yanHΔ* and *yanIΔ* strains. Deletion of *yanH* resulted in a strain where the most prominent compound accumulating is 7-deacetoxyyanuthone A (**3**) (NMR data in Table S6). Interestingly, we also identified two novel compounds in this strain, see Figure 4. Isolation and structure elucidation revealed two C-1 oxidized yanuthone derivatives, which we named yanuthone F (**4**) and yanuthone G (**5**) (NMR data in Tables S7 and S8). Yanuthone G (**5**) is a glycosylated version of yanuthone F (**4**), which can also be detected in trace amounts in the reference strain (see Table S1). Deletion of *yanI* resulted in a strain producing the known compounds 7-deacetoxyyanuthone A (**3**) and 22-deacetylyanuthone A (**6**) (NMR data in Table S9), see Figure 1B. Importantly, the latter compound corresponds to yanuthone E (**2**) without the mevalon moiety. In addition, two novel compounds were produced. The structures were elucidated by NMR spectroscopy revealing that one, which we named yanuthone H (**7**), is very similar to 22-deacetylyanuthone A (**6**), but with a hydroxyl group at C-1 (Figure 4, Table S10). The other novel compound, which we name Yanuthone I (**8**), is a modification of 22-deacetylyanuthone A (**6**) with an oxidative cleavage between C4 and C5 of the sesquiterpene chain (NMR data in Table S11). We note that yanuthone I (**8**) was also detected in trace amounts in the reference strain (see Table S1).

Determination of the yanuthone X₁ structure. As mentioned above yanuthone X₁ (**12**) has an elemental composition corresponding to yanuthone A and C, but was biosynthesized from another precursor than yanuthone D and E. Since the structures of yanuthones A, C, D and E are similar (Figure 1B), it seemed

unlikely that yanuthone X₁ (**12**) was either yanuthone A or C. We therefore isolated and elucidated the structure (see Figure 4 and Table S12). This analysis confirmed that yanuthone X₁ (**12**) does not have the same C₇ core scaffold, but instead a C₆ core with a methoxygroup directly attached to the six-membered ring at the expense of a methyl group (See Figure 4). Despite that yanuthone X₁ (**12**) and yanuthones D and E employ different precursors, they share common features like the epoxide and the sesquiterpene side chain, and we therefore hypothesized that they share common enzymatic steps during their biosynthesis. In agreement with this, examination of the metabolite profiles obtained with the *yan* gene deletion strains revealed that yanuthone X₁ (**12**) was absent in the *yanC*, *yanD*, *yanE*, and *yanG* deletion strains (Table S1). In contrast, yanuthone X₁ (**12**) is produced in larger amounts in the *yanA*Δ strain, which cannot produce 6-MSA.

DISCUSSION

Elucidation of the biosynthetic route from 6-MSA towards yanuthone D. We have used a combination of bioinformatics, genetic tools, chemical analyses and feeding experiments to investigate whether 6-MSA is produced and if it is used for production of toxic secondary metabolites in *A. niger*. Our work demonstrates that 6-MSA is synthesized by the YanA PKS and then subsequently modified into the antimicrobial end product Yanuthone D, which is surprising as yanuthones have previously been suggested to originate from shikimic acid¹⁷. We have also shown that *yanA* defines a gene cluster of ten members, *yanA-I* and *yanR*, which is regulated by YanR, a putative transcription factor with homology to Zn₂Cys₆ TF types that are commonly involved in regulation of secondary metabolite production. The fact that deletion of *yanR* completely abolished production of yanuthone D suggests that YanR acts as an activator of the *yan* cluster. Additionally, analyses of strains where the remaining genes in the *yan* cluster were individually deleted have allowed us to isolate and characterize the full structures of three intermediates. Based on these compounds we propose the entire pathway for yanuthone D formation including addition of a sesquiterpene and a mevalon to the core polyketide moiety at different stages of the biosynthesis (see Figure 6).

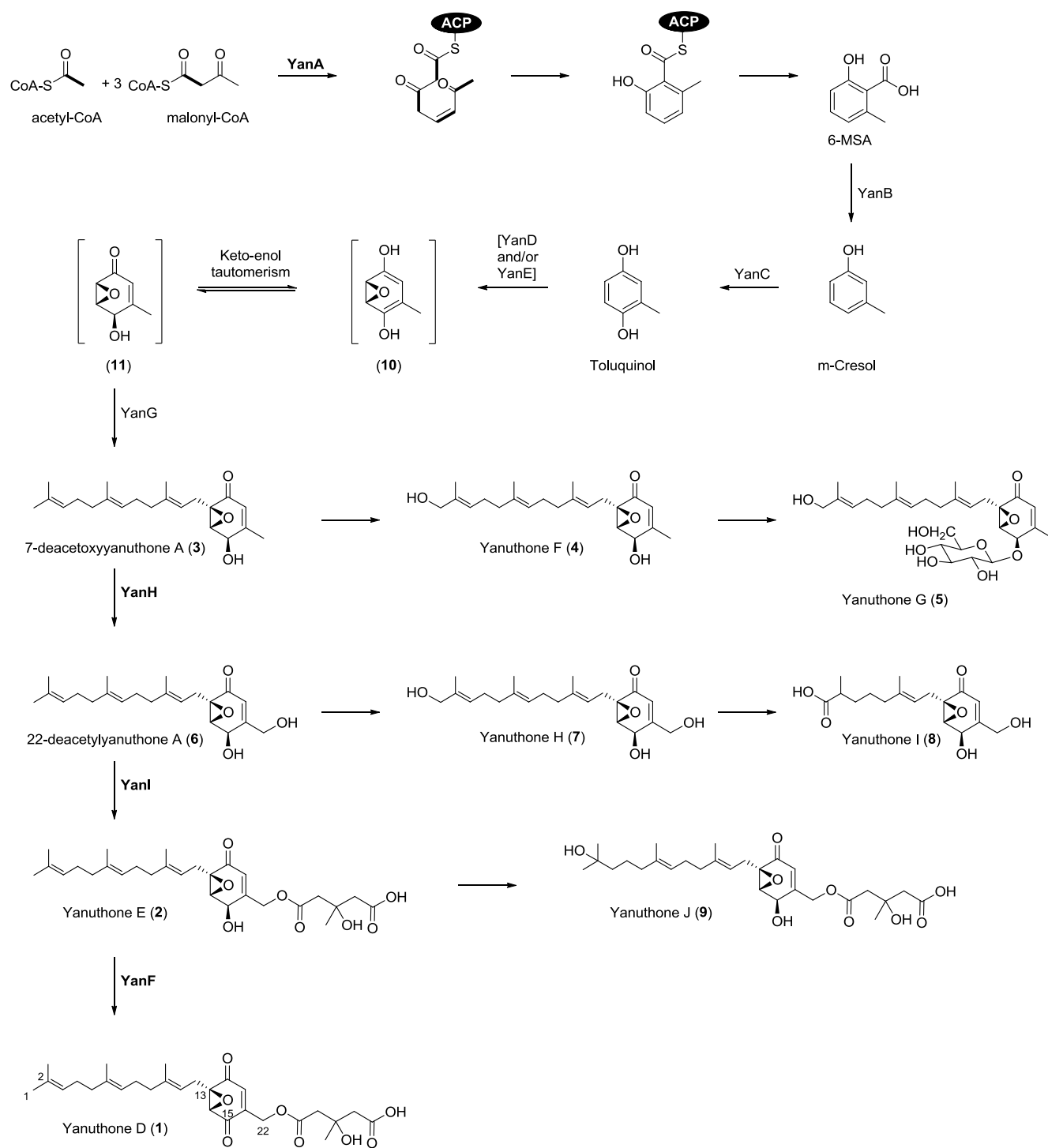


Figure 6. Proposed biosynthesis of yanuthone D (1). Structures and enzymatic activities in brackets are hypothesized, activities in plain text have been proposed from bioinformatics, and activities in bold have been experimentally verified.

In our model, the last intermediate in the pathway is yanuthone E (2), which is converted into the end product yanuthone D (1) by oxidation of the hydroxyl group at C-15 in a process catalysed by YanF. The fact that yanuthone E (2) is present in the reference strain indicates that it may act as a reservoir for rapid conversion into the more potent antibiotic compound yanuthone D. Yanuthone E (2) is likely formed from 22-deacetylanuthone A (6) by attachment of mevalon to the hydroxyl group at C-22. Since 22-deacetylanuthone A (6), but not yanuthone E (2), accumulates in the *yanI*Δ strain we propose that YanI, a putative O-acyl transferase catalyses this step. Intriguingly, YanI therefore appears to be an O-mevalon transferase, an activity, which, to the best of our knowledge, has not previously been described in the literature. Next, we propose that 22-deacetylanuthone A (6) is formed by hydroxylation of C-22 of 7-deacetylanuthone A (3). In agreement with this view, 7-deacetylanuthone A (3), but not 22-deacetoxyyanuthone A (6), accumulates in the absence of YanH.

Unfortunately we did not detect any intermediates leading from 6-MSA to 7-deacetoxyyanuthone A (3) in any of the deletion strains in *A. niger*. The remaining tentative steps in the pathway were therefore deduced from bioinformatics and feeding experiments. First, analyses of patulin formation in *Aspergillus floccosus* (as *A. terreus*, Jens C. Frisvad personal communication) and *A. clavatus* have shown that it requires decarboxylation of 6-MSA into *m*-cresol^{23,24}. This step is catalyzed by 6-MSA decarboxylase²⁵, which has been proposed to be encoded by *patG*²⁴. *m*-cresol is then converted into gentisyl alcohol in two consecutive hydroxylation steps catalyzed by the two P450s, CYP619C3 (PatH) and CYP619C2 (PatI). However, CYP619C2 may also act directly on *m*-cresol to form the co-metabolite toluquinol, which is not an intermediate towards patulin. When we inspected the *yan* gene cluster for similar activities we found a putative 6-MSA decarboxylase (*yanB*) and CYP619C2 (YanC), but not of CYP619C3. These observations suggest that *m*-cresol and toluquinol are intermediates in yanuthone D formation. We present two lines of evidence in support of this view. First, our feeding experiments demonstrate that both compounds can be converted into yanuthone D. Secondly, heterologous expression of *yanA* in *A. nidulans* lead to production of 6-MSA. This compound disappears if the strain also expresses *yanB* indicating that 6-MSA is a substrate for the putative 6-MSA

decarboxylase YanB. Together these results strongly suggest to that *m*-cresol is formed directly from 6-MSA by a decarboxylation reaction, which is most likely catalyzed by YanB. This reaction explains how C₈ based 6-MSA can serve as the building block for the C₇ based core unit of yanuthones. Moreover, the analyses show that toluquinol is an intermediate in the production of yanuthone D and that it is formed from *m*-cresol in a process most likely catalyzed by the putative cytochrome P450 encoded by *yanC*. Conversion of toluquinol into 7-deacetylanuthone A (**3**) requires epoxydation and prenylation. Based on the fact that prenylated toluquinol is never observed in the reference or mutant strains, we propose that epoxydation precedes prenylation. In this scenario, toluquinol is epoxydated into (**10**), which is in equilibrium with the tautomer (**11**). This compound (**11**) is then prenylated to form 7-deacetylanuthone A (**3**) as a sesquiterpene moiety is attached to C-13 of (**11**). The latter reaction is likely catalyzed by YanG, a putative prenyltransferase. This is supported by the observation that yanuthone D (**1**) and all detectable intermediates, including 7-deacetoxyyanuthone A (**3**), were absent in the *yanG*Δ strain. The identity of the gene products(s) responsible for epoxydation of toluquinol is less clear. Amongst the putative activities encoded by the genes in the *yan* cluster, which have not been assigned to any reaction step during the analyses above, we note the presence of a putative dehydrogenase (YanD) and one with an unknown activity (YanE). We hypothesize that one or both of these enzymes catalyze epoxydation. The fact that neither 6-MSA, *m*-cresol, toluquinol, nor any other intermediates were detected in the *yanB*Δ, *yanC*Δ, *yanD*Δ, and *yanE*Δ strains suggests that these small, aromatic compounds must be rapidly degraded, converted into other compound(s), or they may be incorporated into insoluble material e.g. the cell wall.

Accumulation of intermediates in the yanuthone D pathway may trigger detoxification reactions.

Disruption of the biosynthetic pathway towards yanuthone D results in formation of three branch points in the pathway towards yanuthone D: at yanuthone E (**2**), at 7-deacetoxyyanuthone A (**3**), and at 22-deacetylanuthone A (**6**). In addition to yanuthone E (**2**), yanuthone J (**9**) accumulates in the *yanF*Δ strain. Similarly, yanuthone F (**4**) accumulates in addition to 7-deacetoxyyanuthone A (**3**) in the *yanH*Δ strain and Yanuthone H (**7**) accumulates in addition to 22-deacetylanuthone A (**6**) in the *yanI*Δ strain. In all cases, the sesquiterpenes of the

accumulated intermediates in the main pathway are oxidized at C-1 or C-2. As hydroxylation is a known detoxification mode we therefore hypothesize that the abnormally high amount of potentially toxic intermediates 7-deacetoxyyanuthone A (**3**), 22-deacetylyanuthone A (**6**) and yanuthone E (**2**) triggers the cell to initiate phase I type of detoxification processes in which the toxic intermediates are hydroxylated. This hypothesis is supported by the fact that there is no obvious assignment of an enzyme with this activity, encoded by the *yan* gene cluster. An additional variant of yanuthone F (**4**) was identified in the *yanHΔ* strain, in which yanuthone F (**4**) is glycosylated at the hydroxyl group at C-15 to form yanuthone G (**5**). The glucose moiety of yanuthone G (**5**) is intriguing as sugar moieties are rare in fungal secondary metabolites, and the fact that yanuthone G (**5**) is detected in the reference strain shows that it is a naturally occurring compound (see Figure 4 and Table S1). As yanuthone G (**5**) production is up-regulated in *yanHΔ* we suggest that glycosylation poses a second (phase II conjugation) type of mechanism for further detoxification of possible toxic intermediates.

The branch point at 22-deacetylyanuthone A (**6**) revealed a novel compound yanuthone I (**8**), which is identical to 22-deacetylyanuthone A (**6**) and yanuthone H (**7**) but with an oxidative cleavage of the sesquiterpene chain. A similar enzymatic modification has also been proposed for the mycophenolic acid biosynthesis²⁶.

Yanuthone X₁ defines a novel class of yanuthones. Since yanuthones are based on a C₇ scaffold, they were previously proposed to originate from shikimic acid¹⁷. However, in our study we demonstrate that yanuthone D and E origin from the C₈ polyketide precursor 6-MSA, which is decarboxylated to form the C₇ core of the yanuthone structure. In contrast, the novel yanuthone X₁ (**12**) has a C₆ core scaffold that does not origin from 6-MSA and does not require decarboxylation by YanB. Based on this we define two classes of yanuthones, those that are based on the polyketide 6-MSA, class I, and those that are based on the yet unknown precursor leading to the formation of yanuthone X₁ (**12**), class II. The two classes of yanuthones share several enzymatic steps. First we note that the sesquiterpene side chain in yanuthone X₁ (**12**) is likely attached by YanG, as it is the case for yanuthone D. Secondly it depends on enzyme activities of YanC, YanD and YanE, but not of YanB.

Together this suggests that the precursor is a small aromatic compound similar to 6-MSA but lacking the carboxylic acid. Importantly, the main difference between yanuthone D and yanuthone X₁ (**12**) are the groups attached to C-16. In the case of yanuthone X₁ (**12**) this position is oxidized whereas in yanuthones D and E there is a carbon-carbon bond that originates from the methyl group of 6-MSA. Consequently, yanuthone X₁ (**12**) cannot be mevalonated by YanH.

CONCLUSION

This study has identified a cluster of 10 genes, which is responsible for production of antimicrobial yanuthone D in *A. niger*. We show that yanuthone D is based on the polyketide 6-MSA, and not on shikimic acid as previously suggested, and we have proposed a detailed genetic and biochemical pathway for converting 6-MSA into yanuthone D. Interestingly, we have revealed that the novel yanuthone X₁, though similar in structure, is not derived from 6-MSA, but the unknown precursor to yanuthone X₁ does employ several enzymes encoded by the *yan* cluster. An important finding in the elucidation of the biosynthesis is the identification of *yanI* encoding an O-mevalon transferase, which represents a novel enzymatic activity. We have discovered that the pathway towards yanuthone D branches when intermediates accumulate, as three intermediates are hydroxylated. Two of the hydroxylated compounds are further modified by oxidative cleavage of the sesquiterpene and glycosylation, respectively, resulting in five novel yanuthones. The discovery of a glycosylated compound, yanuthone G, is intriguing as glycosylated compounds are very rare in fungal secondary metabolism. With great success, we have employed an interdisciplinary approach for solving the biosynthetic pathway; applying gene deletions, heterologous gene expression, UHPLC-DAD-MS, MS/MS, structural elucidation by NMR spectroscopy and CD, and feeding experiments with labeled and unlabeled metabolites. Together our analyses have cast new insights into understanding the complexity of fungal secondary metabolism.

ACKNOWLEDGMENTS

We kindly acknowledge Kenneth S. Bruno (Pacific Northwest National Laboratory, Richland, WA) for providing the KB1001 strain, and Solveig Kallesøe (University of Copenhagen) for obtaining CD data. We also acknowledge Diana C. Anyaogu for assisting with the RT-qPCR analysis. The study was supported by grant 09-064967 from the Danish Council for Independent Research, Technology, and Production Sciences.

EXPERIMENTAL SECTION

Strains and media. The strain IBT 29539 (*argB2*, *pyrG89*, *veA1*, *nkuAΔ*) was used for strain constructions in *A. nidulans*. ATCC1015-derived strain KB1001 (*pyrGΔ kusA::AFpyrG*) was used for all strain construction in *A. niger*. All fungal strains prepared in the present work (Table S13) have been deposited in the IBT Culture Collection at Department of Systems Biology, DTU, Kgs. Lyngby, Denmark. *Escherichia coli* strain DH5α was used for propagating all plasmids, except *E. coli* ccdB survival2 cells (Invitrogen) were used to propagate plasmids carrying the *ccdB* gene.

Minimal medium for *A. nidulans* was made as described by Cove²⁷ but with 1% glucose, 10 mM NaNO₃ and 2% agar. Minimal medium for *A. niger* was prepared as described by Chiang and co-workers²⁸. YES, MEA and CYA media were prepared as described by Frisvad and Samson²⁹. When necessary, media were supplemented with 4 mM L-arginine, 10 mM uridine, 10 mM uracil and/or 100 μg/ml hygromycin B (Invivogen, San Diego, CA, USA). Luria-Bertani (LB) medium was used for cultivation of *E. coli* strains, and consisted of 10 g/l Tryptone (Bacto), 5 g/l yeast extract (Bacto), and 10 g/l NaCl, pH adjusted to 7.0 prior to autoclavation. When necessary, LB was supplemented with 100 μg/ml ampicillin and/or 25 μg/ml chloramphenicol.

Construction of basic vectors for strain construction. All primers are listed in Table S14. The PfuX7 polymerase³⁰ was used in all PCR reactions. Fragments were assembled via uracil-excision fusion³¹ into a compatible vector.

pDH56 and pDH57 are designed for integration of novel genes into the *IS1* site of *A. nidulans*. In pDH56, the AsiSI/Nb.BtsI uracil-excision cassette of CMBU1111¹⁹ is modified into a *ccdB-cm^R* AsiSI/Nb.BtsI uracil-excision cassette. Unlike with CMBU1111, new fragments can be introduced into this cassette and cloned in *ccdA* deficient *E. coli* strains like DH5α without generating background as false positives resulting from incomplete digestion of the USER cassette are eliminated³². Specifically, the suicide gene *ccdB* and the chloramphenicol resistance gene *cm^R* (*ccdB-cm^R*) construct was PCR amplified (using primers 84 and 85) from

pDONR (Invitrogen) and inserted into CMBU1111 by uracil-excision cloning in a manner that reconstituted the original uracil excision cassette on either side of the insert. Plasmids with the AsiSI/Nb.BtsI uracil-excision cassette are maintained in *ccdB* survival2 cells (Invitrogen). The pDH57 vector was constructed from pDH56 by removing an undesirable Nb.BtsI nicking site located in the *amp^R* gene. pDH56 was PCR amplified in two pieces, using primers 81+76 and 75+ 80. 75 and 76 have complementary uracil-containing overhangs, which were designed to introduce a silent mutation into the Nb.BtsI recognition site. The two PCR fragments were assembled via uracil-excision cloning, and correct clones were verified by sequencing. The gene targeting substrate for insertion of the 6-MSA synthase gene *yanA* was made by amplifying the synthase gene *yanA* (PKS48/ASPNIDRAFT_44965) from IBT29539 genomic DNA by PCR (primers 1+2) and inserted into pDH57, yielding pDH57-*yanA*.

The pDHX2 vector is AMA1-based and designed for episomal gene expression. pDHX2 was constructed by USER fusion of 5 fragments 1) *E. coli* origin of replication (*ori^R*) and the *E. coli* ampicillin resistance gene (*amp^R*) carrying a silent point mutation omitting a Nb.BtsI nicking site, 2) the 5' half of AMA1, 3) the 3' half of AMA1, and 4) 0.5 kb of the *PgpdA* promoter, an AsiSI/Nb.BtsI USER cassette containing *ccdB* and *cm^R*, and the TtrpC terminator, and 5) the *A. fumigatus pyrG* selection marker (Figure S5). Fragment 1 was amplified from pDH57 (primers 77+78), fragment 2, 3, and 5 were amplified from pDEL2 (primers 86+89, 87+88, and 82+83)³³, fragment 4 was amplified from pDH57 (primers 79+80). Fragments were assembled as described by Geu-Flores and coworkers³¹ using equal molar amounts of purified PCR product, and correct clones were verified by restriction digestion. Plasmids for episomal heterologous expression of cluster genes were constructed by PCR amplification of ORFs using primers 3-12 pairwise. Genes were inserted into AsiSI/Nb.BtsI digested pDHX2 as described by Nour-Eldin and co-workers³⁴, resulting in pDHX2-*yanB*, pDHX2-*yanC*, pDHX2-*yanD*, and pDHX2-*yanE*. Plasmids were verified by sequencing.

Plasmids carrying gene targeting substrates for gene deletion in *A. niger* were constructed by PCR amplification of upstream (US) and downstream (DS) targeting sequences along with the *hph* marker, conferring resistance to hygromycin B. US and DS targeting sequences were generated using the primers 17-72,

hph was amplified from pCB1003³⁵ using primers 13+14. The three fragments were assembled into the CMBU0020 vector¹⁹.

Strain construction. Protoplasting and gene-targeting procedures were performed as described previously for *A. nidulans*^{36,37} and *A. niger*²⁸. Gene targeting substrate for insertion of the 6-MSA synthase (*yanA*) was excised from pDH57-*yanA* by NotI digestion and transformed into IBT29539. Transformants, with *yanA* integrated into *IS1*, were verified by diagnostic PCR as described by Hansen and co-workers¹⁹.

Strains for episomal expression of cluster genes were constructed by transforming the IS1-*yanA* strain with circular plasmids (pDHX2-*yanB*, pDHX2-*yanC*, pDHX2-*yanD*, and pDHX2-*yanE*) using *pyrG* as a selectable marker.

A. niger deletion strains were constructed by transforming KB1001 with bipartite gene targeting substrates. The substrates were generated by PCR amplification of the US::*hph*::DS cassettes of the CMBU0020-based plasmids using primers GENE_US-FW+73 and 74+GENE_DS-RV. Deletion strains were selected on 100 µg/ml hygromycin B, and verified by diagnostic PCR.

RNA extraction and RT-qPCR. RNA isolation from the *A. niger* strains and subsequent quantitative RT-PCR reactions were done as previously described by Hansen and co-workers¹⁹ except that biomass for RNA isolation was prepared with a Tissue-Lyser LT (Qiagen) by treating samples for 1 min at 45 MHz. The *A. niger* histone 3 encoding gene, *hhtA* (ASPNIIDRAFT_52637) and gamma-actin encoding gene *actA* (ASPNIIDRAFT_200483) were used as internal standards for normalization of expression levels. All primers used for quantitative RT-PCR are shown in Table S14 (primers 90-121).

Chemical analysis of strains. Unless otherwise stated strains were inoculated as a three point inoculation on solid MM media and incubated at 37 °C for 5 days. Extraction of metabolites was performed as described by Smedsgaard³⁸. 6-MSA was purchased from (Apin Chemicals, Oxon, UK). Analysis was performed using ultra-

high-performance liquid chromatography (UHPLC) UV/Vis diode array detector (DAD) high-resolution MS (TOFMS) on a maXis 3G orthogonal acceleration quadrupole time of flight mass spectrometer (Bruker Daltonics, Bremen, Germany) equipped with an electrospray ionization (ESI) source and connected to an Ultimate 3000 UHPLC system (Dionex, Sunnyvale, USA). The column used was a reverse-phase Kinetex 2.6- μm C₁₈, 100 mm \times 2.1mm (Phenomenex, Torrance, CA), and the column temperature was maintained at 40 °C. A linear water- acetonitrile (LCMS-grade) gradient was used (both solvents were buffered with 20 mM formic acid) starting from 10% (vol/vol) acetonitrile and increased to 100% in 10 min, maintaining this rate for 3 min before returning to the starting conditions in 0.1 min and staying there for 2.4 min before the following run. A flow rate of 0.4 mL·min⁻¹ was used. TOFMS was performed in ESI⁺ with a data acquisition range of 10 scans per second at m/z 100–1000. The TOFMS was calibrated using Bruker Daltonics high precision calibration algorithm (HPC) by means of the use of the internal standard sodium formate, which was automatically infused before each run. This provided a mass accuracy of better than 1.5 ppm in MS mode. UV/VIS spectra were collected at wavelengths from 200 to 700 nm. Data processing was performed using DataAnalysis 4.0 and Target Analysis 1.2 software (Bruker Daltonics). Tandem MS was performed with fragmentation energy was varied from 18 to 55 eV.

Preparative isolation of selected metabolites. The fungal strains was inoculated as three point inoculations on 10-200 plates of YES medium and incubated at 30 °C for 5 days. The plates were harvested and extracted twice overnight with ethyl acetate (EtOAc) with or without 1 % formic acid. For specific details about each extract see supplementary Table S2. The extracts were filtered and concentrated in vacuo. The combined extract was dissolved in methanol (MeOH) and H₂O (9:1), and an equal amount of heptane was added where after the phases were separated. To the MeOH/H₂O phase H₂O was added to a ratio of 1:1, and metabolites were then extracted with dichlormethane (DCM). The phases were then concentrated separately in vacuo. The DCM phase was adsorbed onto diol column material and dried before packing into a SNAP column (Biotage, Uppsala, Sweden) with diol material. The extract was then fractionated on an Isolera flash purification system (Biotage) using

seven steps of heptane-DCM-EtOAc-MeOH. Solvents used were of HPLC grade and H₂O was purified and deionized by a Millipore system through a 0.22 µm membrane filter (milliQ-water). For flow rate, column sizes and fractionation see Table S2.

The Isolera fractions were subjected to further purification on a semi-preparative HPLC which was either a Waters 600 Controller with a 996 photodiode array detector (Waters, Milford, MA, USA) or a Gilson 322 Controller connected to a 215 Liquid Handler, 819 Injection Module and a 172 DAD (Gilson, Middleton, WI, USA). This was achieved using a Luna II C₁₈ column (250 x 10 mm, 5 µm, Phenomenex) or a Gemini C6-Phenyl 110A column (250 x 10.00 mm, 5 µm, Phenomenex). 50 ppm TFA was added to acetonitrile of HPLC grade and milliQ-water. For choice of system, flow rate, column, gradients and yields see Table S2.

NMR and structural elucidation. The 1D and 2D spectra were recorded on a Unity Inova-500 MHz spectrometer (Varian, Palo Alto, CA, USA). Spectra were acquired using standard pulse sequences and ¹H spectra as well as DQF-COSY, NOESY, HSQC and HMBC spectra were acquired. The deuterated solvent was acetonitrile-*d*₃ and signals were referenced by solvent signals for acetonitrile-*d*₃ at δ_H = 1.94 ppm and δ_C = 1.32/118.26 ppm. The NMR data was processed in Bruker Topspin 3.1 or ACD NMR Workbook. Chemical shifts are reported in ppm (δ) and scalar couplings in hertz (Hz). The sizes of the *J* coupling constants reported in the tables are the experimentally measured values from the spectra. There are minor variations in the measurements which may be explained by the uncertainty of *J*. The description of the structural elucidation of the yanuthones can be found in the supplementary material and NMR data for all compounds are listed in Tables S3-S12.

Circular dichroism (CD) spectra were obtained using a J-710 spectropolarimeter (Jasco, Cremella, Italy). The samples were dissolved in methanol to a concentration of 1 mg/3 mL and analysed in 0.2 cm optical path length cells. The temperature was 20 °C and the spectra were recorded from 200 – 500 nm. Optical rotation was measured on a Perkin Elmer 321 Polarimeter.

Production and purification of fully labeled ^{13}C -6-MSA. As fully labeled $^{13}\text{C}_8$ -6-MSA was not commercially available it was produced in-house from the 6-MSA producing strain by batch cultivation using a Sartorius 1L bioreactor (Sartorius, Stedim Biotech, Germany) with a working volume of 0.8 L equipped with 2 Rushton six-blade disc turbines. The pH electrode (Mettler Toledo, OH, USA) was calibrated according to manufacture standard procedures. The bioreactor was sparged with sterile atmospheric air and off-gas concentrations of oxygen and carbon dioxide were measured with a Prima Pro Process Mass Spectrometer (Thermo-Fischer Scientific, MA). Temperature was maintained at 30°C throughout the cultivation and pH controlled by automatic addition of 2M NaOH and H_2SO_4 . Start culture conditions were pH: 3.0, stir rate: 100 rpm and air flow: 0.1 volume of air per volume of liquid per minute (vvm). Following inoculation these conditions were changed in a linear fashion over a period of 720 min. to pH: 5.0, stir rate: 800 rpm and air flow: 1 vvm. The strain was cultivated until glucose was depleted, measured by glucose test strips (Machery-Nagel, Germany) and the culture had entered stationary phase as monitored by off-gas CO_2 concentration.

When glucose was depleted the entire volume of the reactor was harvested and the biomass removed by filtration through a Whatman 1 qualitative paper filter followed by centrifugation at 8000 g for 20 min. to remove fine sediments. The 6-MSA was then recovered from the supernatant by liquid-liquid extraction using ethyl acetate containing 0.5 % formic acid.

The organic extract then completely dried in vacuo to give a crude extract that was redissolved in 20 mL ethyl acetate and dry loaded onto 3 g Septra ZT C18 (Phenomenex) resin prior packing into a 25 g SNAP column (Biotage, Uppsala, Sweden) with 22 g pure resin in the base. The crude extract was then fractionated on an Isolera flask purification system (Biotage, Uppsala, Sweden) using an water-acetonitrile gradient starting at 15:85 going to 100 % acetonitrile in 23 min at a flow rate of 25 mL min^{-1} and kept at that level for an additional 4 min. Fractions were automatically collected using UV detection at 210 and 254 nm resulting in a total of 15 fractions of which 3 were pooled and analysed. 6-MSA concentration was assessed using a Dionex Ultimate 3000 UHPLC coupled with a ultimate 3000 RS diode array detector (Dionex, Germering, Germany) equipped

with a Poroshell 120 phenyl hexyl 2.1x100 mm, 2.7 μ m (Agilent, Little River, DE) column. Finally, purity (98.7 %) was analyzed by UHPLC-TOFMS (Figure S3A).

Feeding experiments. Solid YES plates were prepared for each strain using a 6 mm plug drill to make a well in the middle of the agar. The well should act as a feeding container. Water, 25-100 μ l, was added to the well and spores of the appropriate strains were inoculated into the water and the plates were incubated at 30 °C for 5 days. 100 μ g 13 C-6-MSA, *m*-cresol, toluquinol (=ortoluquinol) (in 25-100 μ l 1:4 ethanol:water) was added to the three plates after 24h, 48h and 72h of incubation, respectively. Control plates without feed were included. Agar plugs were taken both as reported previously³⁸, and also separately from the center of the colony, the middle ring of the colony, and the rim of the colony, respectively, to verify diffusion and absorption of the 6-MSA and the location of yanuthone production. Metabolites were analyzed as described in the section “Chemical analysis of strains”.

REFERENCES:

- (1) Williams, D.H.; Stone, M.J.; Hauck, P.R.; Rahman, S.K.; *J. Nat. Prod.* **1989**, *52*, 1189-1208.
- (2) Olsen, J.H.; Dragstead, L.; Autrup, H.; *Brit. J. Cancer* **1988**, *58*, 392-396.
- (3) Endo, A.; Kuroda, M.; Tsujita Y.; *J. Antibiot.* **1976**, *29*, 12, 1346-1348.
- (4) Bentley, R.; *Chem. Rev.* **2000**, *100*, 3801-3825.
- (5) Frisvad, J.C.; Rank, C.; Nielsen, K.F.; Larsen, T.O.; *Med. Mycol.* **2009**, *47*, S71.
- (6) Rho, M.C.; Toyoshima, M.; Hayashi, M.; Uchida, R.; Shiomi, K.; Komiyama, K.; Omura, S.; *J. Antibiot.* **1998**, *51* (1), 68-72.
- (7) Pel, H.J.; de Winde, J.H.; Archer, D.B.; Dyer, P.S.; Hofmann, G.; Schaap, P.J.; Turner, G.; de Vries, R.P.; Albang, R.; Albermann, K.; Andersen, M.R.; Bendtsen, J.D.; Benen, J.A E.; van den Berg, M.; Breestraat, S.; Caddick, M.X.; Contreras, R.; Cornell, M.; Coutinho, P.M.; Danchin, E.G.J.; Debets, A.J.M.; Dekker, P.; van Dijk, P.W.M.; van Dijk, A.; Dijkhuizen, L.; Driessen, A.J.M.; d'Enfert, C.; Geysens, S.; Goosen, C.; Groot, G.S.P.; de Groot, P.W.J.; • Guillemette, T.; Henrissat, B.; Herweijer, M.; van den Hombergh, J.P.T.W.; van den Hondel, C.A.M.J.J.; van der Heijden, R.T.J.M.; van der Kaaij, R.M.; Klis, F.M.; Kools, H.J.; Kubicek, C.P.; van Kuyk, P.A.; Lauber, J.; Lu, X.; van der Maarel, M.J.E.C.; Meulenberg, R.; Menke, H.; Mortimer, M.A.; Nielsen, J.; Oliver, S.G.; Olsthoorn, M.; Pal, K.; van Peij, N.N.M.E.; Ram, A.F.J.; Rinas, U.; Roubos, J.A.; Sagt, C.M.J.; Schmoll, M.; Sun, J.; Ussery, D.; Varga, J.; Vervecken, W.; van de Vondervoort, P.J.J.; Wedler, H.; Wösten, H.A.B.; Zeng, A.; van Ooyen, A.J.J.; Visser, J.; Stam, H.; *Nat. Biotechnol.* **2007**, *25*, 2:221-231.
- (8) Andersen, M.R.; Nielsen, J.B.; Klitgaard, A.; Petersen, L.M.; Zachariasen, M.; Hansen, T.J.; Blicher, L.H.; Gotfredsen, C.H.; Larsen, T.O.; Nielsen, K.F.; Mortensen, U.H. *PNAS* **2013**, *110*, E99–107.
- (9) Gross, H.; *Appl. Microbiol. Biotechnol.* **2007**, *75*, 267-277.
- (10) Cox, R. J. *Organic & biomolecular chemistry* **2007**, *5*, 2010–2026.
- (11) Hertweck, C. *Angewandte Chemie (International ed. in English)* **2009**, *48*, 4688–4716.
- (12) Baker, S.E.; *Med. Mycol.* **2006**, *44*, S17–21.
- (13) Frisvad, J.C.; Smedsgaard, J.; Samson, R.A.; Larsen, T.O.; Thrane, U.; *J. Agr. Food Chem.* **2007**, *55*, 9727–9732.
- (14) Fisch, K.M.; Gillaspay, A.F.; Gipson, M.; Henrikson, J.C.; Hoover, A.R.; Jackson, L.; Najar, F.Z.; Wägele, H.; Cichewicz, R.H.; *J. Ind. Microbiol. Biotechnol.* **2009**, *36*, 1199–1213.
- (15) Wattanachaisaereekul, S.; Lantz, A.E.; Nielsen, M.L.; Nielsen, J.; *Metab. Eng.* **2008**, *10*, 246– 254.
- (16) Beck, J.; Ripka, S.; Siegner, A.; Schiltz, E.; Schweizer, E.; *Eur. J. Biochem.* **1990**, *192*, 487-498.
- (17) Bugni, T.S.; Abbanat, D.; Bernan, V.S.; Maiese, W.M.; Greenstein, M.; Van Wagoner, R.M.; Ireland, C.M.; *J. Org. Chem.* **2000**, *65*, 7195-7200.

- (18) Li, X.; Choi, H.D.; Kang, J.S.; Lee, C.O.; Son, B.W.; *J. Nat. Prod.* **2003**, *66*, 1499–1500.
- (19) Hansen, B.G.; Salomonsen, B.; Nielsen, M.T.; Nielsen, J.B.; Hansen, N.B.; Nielsen, K.F.; Regueira, T.B.; Nielsen, J.; Patil, K.R.; Mortensen, U.H.; *Appl. Environ. Microbiol.* **2011**, *77*, 3044–3051.
- (20) Nielsen, M.L.; Nielsen, J.B.; Rank, C.; Klejnstrup, M.L.; Holm, D.K.; Brogaard, K.H.; Hansen, B.G.; Frisvad, J.C.; Larsen, T.O.; Mortensen, U.H.; *FEMS Microbiol. Lett.* **2011**, *321*, 2:157–166.
- (21) Stanke, M.; Morgenstern, B.; *Nucleic Acids Res.* **2005**, *33* (Web Server issue):W465–W467.
- (22) Marchler-Bauer, A.; Lu, S.; Anderson, J.B.; Chitsaz, F.; Derbyshire, M.K.; DeWeese-Scott, C.; Fong, J.H.; Geer, L.Y.; Geer, R.C.; Gonzales, N.R.; Gwadz, M.; Hurwitz, D.I.; Jackson, J.D.; Ke, Z.; Lanczycki, C.J.; Lu, F.; Marchler, G.H.; Mullokandov, M.; Omelchenko, M.V.; Robertson, C.L.; Song, J.S.; Thanki, N.; Yamashita, R.A.; Zhang, D.; Zhang, N.; Zheng, C.; Bryant, S.H.; *Nucleic Acids Res.* **2011**, *39*, 225–229.
- (23) Artigot, M.P.; Loiseau, N.; Laffitte, J.; Mas-Reguieg, L.; Tadrist, S.; Oswald, I.P.; Puel, O.; *Microbiology* **2009**, *155*, 1738–1747.
- (24) Puel, O.; Galtier, P.; Oswald, I.P.; *Toxins* **2010**, *2*, 613–631.
- (25) Light, R.J.; *Biochim. Biophys. Acta* **1969**, *191*, 2:430–438.
- (26) Regueira, T.B.; Kildegaard, K.R.; Hansen, B.G.; Mortensen, U.H.; Hertweck, C.; Nielsen, J.; *Appl. Environ. Microbiol.* **2011**, *77*, 9: 3035–3043.
- (27) Cove, D. J.; *Biochim. Biophys. Acta.* **1966**, *113*, 51–56.
- (28) Chiang, Y.-M.; Meyer, K.M.; Praseuth, M.; Baker, S.E.; Bruno, K.S.; Wang, C.C.C.; *Fungal Genet. Biol.* **2011**, *48*, 430–437.
- (29) Samson, R.A.; Houbaken, J.; Thrane, U.; Frisvad, J.C.; Andersen, B.. Food and Indoor Fungi. CBS Laboratory Manual Series 2, CBS KNAW Fungal Biodiversity Centre: Utrecht (NL), 2010.
- (30) Nørholm, M.H.H.; *BMC Biotechnol.* **2010**, *10*, 21.
- (31) Geu-Flores, F.; Nour-Eldin, H.H.; Nielsen, M.T.; Halkier, B.A.; *Nucleic Acids Res.* **2007**, *35*, 7, e55.
- (32) Bernard, P.; Couturier, M.; *J Mol. Biol.* **1992**, *226*(3), 735–745.
- (33) Nielsen, J.B.; Nielsen, M.L.; Mortensen, U.H.; *Fungal Genet. Biol.* **2008**, *45*, 165–170.
- (34) Nour-Eldin, H.H.; Hansen, B.G.; Nørholm, M.H.H.; Jensen, J.K. and Halkier, B.A.; *Nucleic Acids Res.* **2006**, *34*, 18, e122.
- (35) McCluskey, K.; Wiest, A.; Plamann, M.; *J. Biosci.* **2010**, *35*: 119–126.
- (36) Johnstone, I.L.; Hughes, S.G.; Clutterbuck, A.J.; *EMBO J.* **1985**, *4*, 1307–1311.

(37) Nielsen, M.L.; Albertsen, L.; Lettier, G.; Nielsen, J.B.; Mortensen, U.H.; *Fungal Genet. Biol.* **2006**, *43*, 54-64.

(38) Smedsgaard, J.; *J. Chromatogr. A* **1997**, *760*, 264-270.

Chapter 11

Heterologous Reconstitution of the Intact Geodin Gene Cluster in *Aspergillus nidulans* Through a Simple and Versatile PCR Based Approach

This chapter contains the paper "Heterologous Reconstitution of the Intact Geodin Gene Cluster in *Aspergillus nidulans* Through a Simple and Versatile PCR Based Approach", published in *PLoS ONE*, of which I am co-author.

In this study we have used *A. nidulans* as a host for heterologous expression of the entire (+)-geodin cluster from *A. terreus*. After successful transfer to *A. nidulans*, selected steps in the pathway were investigated using gene deletion and gene expression approaches.

My contribution to this study was development of Integration Site 3 (IS3), and preparation of strains for testing autoregulation of the transcriptional regulator. Supporting information is included in Appendix F.

Heterologous Reconstitution of the Intact Geodin Gene Cluster in *Aspergillus nidulans* through a Simple and Versatile PCR Based Approach

Morten Thrane Nielsen^{1*}, Jakob Blæsbjerg Nielsen², Dianna Chinyere Anyaogu^{1*}, Dorte Koefoed Holm, Kristian Fog Nielsen, Thomas Ostenfeld Larsen^{1*}, Uffe Hasbro Mortensen^{1*}

Department of Systems Biology, Technical University of Denmark, Kgs. Lyngby, Denmark

Abstract

Fungal natural products are a rich resource for bioactive molecules. To fully exploit this potential it is necessary to link genes to metabolites. Genetic information for numerous putative biosynthetic pathways has become available in recent years through genome sequencing. However, the lack of solid methodology for genetic manipulation of most species severely hampers pathway characterization. Here we present a simple PCR based approach for heterologous reconstitution of intact gene clusters. Specifically, the putative gene cluster responsible for geodin production from *Aspergillus terreus* was transferred in a two step procedure to an expression platform in *A. nidulans*. The individual cluster fragments were generated by PCR and assembled via efficient USER fusion prior to transformation and integration via re-iterative gene targeting. A total of 13 open reading frames contained in 25 kb of DNA were successfully transferred between the two species enabling geodin synthesis in *A. nidulans*. Subsequently, functions of three genes in the cluster were validated by genetic and chemical analyses. Specifically, ATEG_08451 (*gedC*) encodes a polyketide synthase, ATEG_08453 (*gedR*) encodes a transcription factor responsible for activation of the geodin gene cluster and ATEG_08460 (*gedL*) encodes a halogenase that catalyzes conversion of sulochrin to dihydrogeodin. We expect that our approach for transferring intact biosynthetic pathways to a fungus with a well developed genetic toolbox will be instrumental in characterizing the many exciting pathways for secondary metabolite production that are currently being uncovered by the fungal genome sequencing projects.

Citation: Nielsen MT, Nielsen JB, Anyaogu DC, Holm DK, Nielsen KF, et al. (2013) Heterologous Reconstitution of the Intact Geodin Gene Cluster in *Aspergillus nidulans* through a Simple and Versatile PCR Based Approach. PLoS ONE 8(8): e72871. doi:10.1371/journal.pone.0072871

Editor: Marie-Joelle Vioille, University Paris South, France

Received: November 15, 2012; **Accepted:** July 19, 2013; **Published:** August 23, 2013

Copyright: © 2013 Nielsen et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: This work was supported by the Danish Research Agency for Technology and Production, grant 09-064967. The PhD studies of which this work was part of was funded by the Research School for Biotechnology at the faculty of Life Sciences, University of Copenhagen. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: tol@bio.dtu.dk (TOL); um@bio.dtu.dk (UHM)

† Current address: Novo Nordisk Foundation Center for Biosustainability, Technical University of Denmark, Hørsholm, Denmark

‡ These authors contributed equally to this work.

Introduction

Fungal natural products constitute a rich resource for bioactive secondary metabolites [1,2]. To fully exploit this potential, it is essential to identify the genes required for the biosynthesis of these compounds. This process is becoming progressively easier due to the rapidly increasing number of fungal genomes that have been fully sequenced; and since the genes involved in the production of a given secondary metabolite often cluster together in the same region of a chromosome [1–3]. Importantly, the genome sequencing projects have revealed that the number of putative gene clusters for secondary metabolite production greatly exceeds the number of known natural products in a given fungus, hence, indicating that most fungal compounds are yet to be discovered.

The prerequisite for genetic exploration of the huge reservoir of undiscovered biosynthetic pathways is solid methodologies for cultivating, propagating and genetically manipulating the producing species. However, the vast majority of newly sequenced organisms fail to meet these requirements, hence, hampering

pathway elucidation and exploitation. An attractive solution to this problem is to transfer pathways into another fungus where the methodology is well established. This approach has been used successfully to investigate several gene clusters [4–6]. All of these studies apply a strategy where individual genes in a cluster are PCR amplified, cloned, and integrated sequentially into either a random or defined locus. One advantage of this strategy is that it allows easy engineering of individual genes prior to integration in the host strain. In addition, it is possible to insert the foreign gene(s) into a well characterized locus that supports high expression levels [7]. Inserting the genes into a known locus also simplifies strain validation. In a recent example of this strategy, Itoh et al. [6] transferred five genes from *Aspergillus fumigatus* to *A. oryzae* allowing the authors to deduce the biosynthetic route for the meroterpenoid pyropyripene A. However, the strategy may be limited to reconstitution of simple pathways depending on only a small number of genes since assembly of multistep pathways will require several rounds of tedious iterative integration steps or enough genetic markers.

For multistep pathways it is therefore desirable to transfer entire gene clusters from the natural producer to the expression host in one or a few steps. This requires a host, which can efficiently express all genes in the cluster and correctly splice the resulting transcripts. Assembly of multiple genes and PCR fragments can efficiently be performed in *Saccharomyces cerevisiae* by recombination based methods [8]. However, the demand for splicing disfavors *S. cerevisiae* as expression host for this gene transfer strategy as only little splicing occurs in yeast. In contrast, correct and efficient splicing of heterologous transcripts by another filamentous fungus is likely. Activation of the cluster in the new host requires as a minimum that the chromatin structure at the insertion site is in an open configuration and that transcription factors exist that recognize the individual promoters in the cluster. The latter may be facilitated by the fact that many clusters appear to contain one or more genes that encode transcription factors. In a pioneering study, Bergmann et al. showed that expression of such a gene activated the entire aspyridone gene cluster in *A. nidulans* [9]. The potential of transferring entire clusters to an expression host has been demonstrated by Sakai et al. who managed to produce citrinin in *A. oryzae* [10]. To obtain this feat, they isolated and transformed a cosmid from a *Monascus purpureus* library containing all six genes required for production of citrinin into this production host. Here constitutive expression of the citrinin pathway regulator encoded by *ctn4* dramatically increased citrinin production in the heterologous host. However, construction and screening of cosmid libraries is not simple and a versatile PCR based method that facilitates the transfer of entire gene clusters from the natural producer to the expression host is desirable.

We have previously developed a versatile PCR based expression platform that can be used for heterologous expression in *A. nidulans* of one or a pair of genes from the defined locus *ISI*, which supports a high level of gene expression [4,7]. Here, we demonstrate how this platform can be expanded to allow transfer of an entire gene cluster from another fungus into *ISI*.

Results and Discussion

2.1 Method for PCR based reconstruction of fungal gene clusters in a heterologous host

Our method for transfer of large DNA fragments relies on successive gene targeting events that introduce ~15 kb fragments into a defined locus. In the example of the method presented here, we transfer a gene cluster into *ISI* taking advantage of a vector set we have previously developed for this purpose [7]. In our method, fragments covering the entire gene cluster are PCR amplified, combined via USER fusion into ~15 kb fragments, and inserted into the integration vector by USER cloning (Figure 1A) [11,12]. The first fragment to be integrated is assembled in the vector and integrated into *ISI* as we have described previously for integration of single genes [7]. The following fragments are integrated as an extension of the previous one by using one of two different markers, *argB* and *pyrG*, for selection. In principle, an indefinite number of integrations can be done, since the marker from the previous integration is excised as the new fragment with the other marker integrates (Figure 1B). This principle is referred to as re-iterative gene targeting [13]. Importantly, marker replacements allows for a simple selection scheme for identification of correctly targeted strains. If, as in our case, the *pyrG* marker is flanked by a direct repeat, we recommend to use the *pyrG* marker in the last integration step, as the *pyrG* marker subsequently can be removed by direct repeat recombination if desirable [14]. In this manner both markers are available in the finalized strain providing a marker repertoire for additional genetic engineering.

In the present study we demonstrate the potential of our method by transferring the geodin gene cluster from *A. terreus* to *A. nidulans*. This cluster was chosen, firstly, because it contains a gene encoding a putative transcription factor, which potentially could facilitate activation of the other genes in the cluster. Secondly, the biosynthetic pathway for geodin production is partially characterized (Figure 2) [15–21], which simplifies the delineation of the cluster size. Thirdly, the geodin pathway shares several steps with the monodictyphenone pathway including production of several common intermediates/products e.g. emodin [22]. We therefore envisioned, that the chance of producing geodin in *A. nidulans* would be increased, as the corresponding endogenous enzymes could complement geodin enzymes that might not be functional. Moreover, shared intermediates would likely be non-toxic to the host.

2.2 Delineation of the putative geodin producing gene cluster in *A. terreus*

Three enzymes involved in geodin production have previously been linked to genes. Specifically, dihydrogeodin oxidase, a polyketide synthase, and a thioesterase are encoded by ATEG_08458 [18], ATEG_08451 [19–21], and ATEG_08450 [20], respectively, see *Aspergillus* Comparative Sequencing Project database (Broad Institute of Harvard and MIT, <http://www.broadinstitute.org/>). As indicated by the gene numbers these genes localize in close proximity to each other strongly suggesting that a gene cluster responsible for production of geodin exists. Three additional enzymes required for geodin biosynthesis have been characterized biochemically *in vitro*: emodin anthrone oxygenase [15], emodin O-methyltransferase [16] and questin oxygenase [17]. Moreover, the occurrence of chlorine atoms in geodin suggests the involvement of a halogenase.

To explore the possibility that genes encoding these four enzymatic activities were also present in this region, we examined all annotated open reading frames (ORFs) positioned between ATEG_08458 and ATEG_08450, as well as 20 kb upstream of ATEG_08458 and downstream of ATEG_08450. Among the ORFs in this region, none had a functional annotation corresponding to these enzymatic activities. We therefore subjected all annotated ORFs in this regions to a functional prediction using the BLAST algorithm [23] from NCBI and the HHpred software [24]. This analysis uncovered three genes that could encode putative methyltransferases (ATEG_08449, ATEG_08452 and ATEG_08456), one ORF that may encode an oxygenase carrying out a Baeyer-Villiger oxidation (ATEG_08459), and a putative halogenase (ATEG_08460), see Table 1.

Unexpectedly, none of the annotated ORFs were found to encode the emodin anthrone oxygenase (Figure 2). To investigate this apparent dilemma, we searched the literature for other oxygenases catalyzing a similar reaction. Via this effort, we found an oxygenase that catalyzes conversion of norsolorinic acid anthrone to norsolorinic acid, a step towards aflatoxin production in *A. flavus* [25]. This recently identified enzyme is encoded by the gene *hypC*. Inspired by these findings, we used the sequence of *HypC* to conduct pair-wise alignments to putative proteins encoded by alternative ORFs in the proposed geodin gene cluster. One short putative ORF encodes a protein of 150 amino acid residues with an overall identity of 34% with the 210 residues of *HypC*. Moreover, the conserved amino acid residues were primarily positioned in catalytic regions or conserved domains (Figure S1, [25]). Interestingly, the putative ORF is oriented in the opposite direction of ATEG_08457. This strongly indicates that the region at ATEG_08457 is wrongly annotated and contains two separate ORFs that we now denote ATEG_08457-1 (the originally

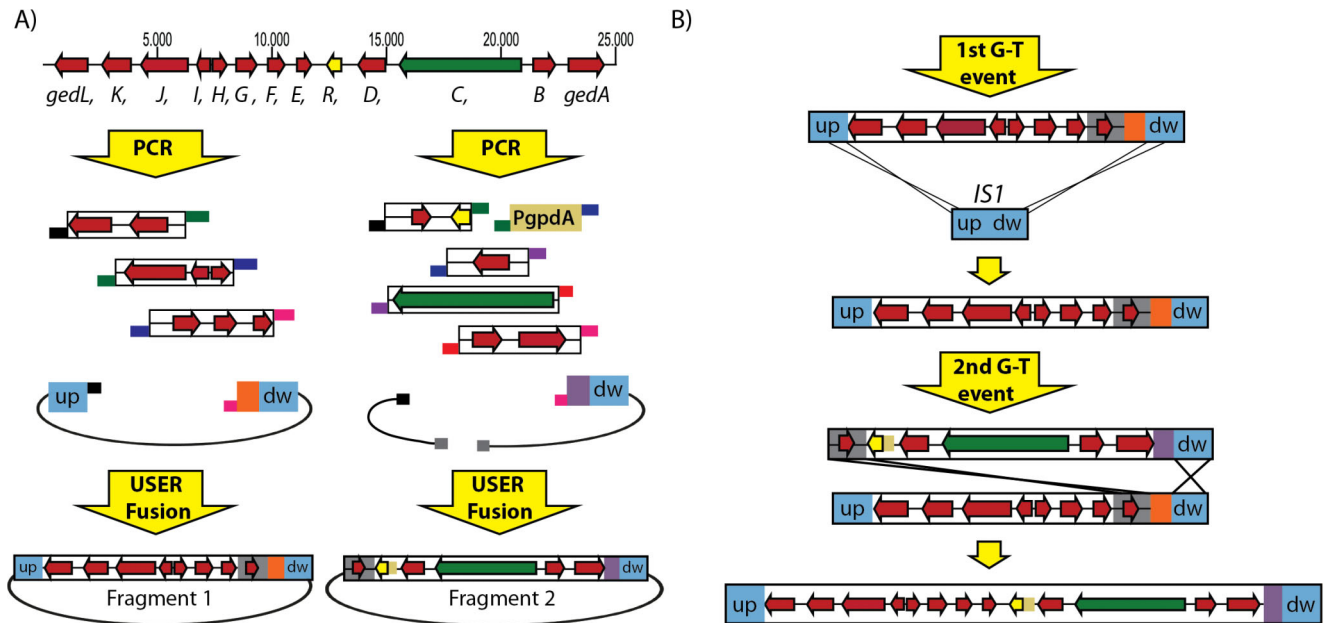


Figure 1. Schematic overview of the PCR based USER cloning strategy for transfer of entire gene clusters from one fungus to another. In the illustrated case, the geodin gene cluster in *A. terreus* is PCR amplified, cloned, and integrated into the *IS1* locus in *A. nidulans*. A) ORFs GedA-GedL are depicted as arrows. The yellow and green arrows represent the ORFs encoding the transcription factor and the PKS, respectively. Remaining ORFs are represented by red arrows. Arrow size is proportional to ORF length and arrow direction indicates genomic orientation. Numbers above the gene cluster specify sequence in base pairs. Genomic DNA fragments and cloning vectors are amplified as PCR products using primers extended with uracil-containing tails. The tails contain matching sequences (indicated by identical colors) allowing for PCR product assembly in a single USER Fusion reaction. For the geodin cluster, all putative ORFs are fused into two fragments, which are individually inserted into a vector prepared for gene targeting. Blue boxes labeled up (upstream) and dw (downstream) represent targeting sequences for homologous recombination into *IS1* in the first gene-targeting event. The targeting sequences in the second integration event are represented in gray and blue and consist of the overlapping region between Fragment 1 and 2 and the downstream part of *IS1*, respectively. Genetic markers used for selection are depicted in orange (*argB*) and purple (*AFpyrG*). The sizes of uracil-containing tails, vector elements and PgpA fragment are not drawn to scale. B) The first gene-targeting event introduces the first fragment into *IS1* by homologous recombination between *IS1* up and down-sequences as indicated. The second gene-targeting event introduces the second fragment using the overlapping region of the Fragment 1 and 2 (gray) and the downstream section of *IS1* as targeting sequences. Note that additional DNA can be inserted in subsequent gene-targeting events. For example, a third fragment can be inserted by using the downstream end of fragment 2 and the downstream region of *IS1* as targeting sequences. See text for details concerning use and recycling of markers.

doi:10.1371/journal.pone.0072871.g001

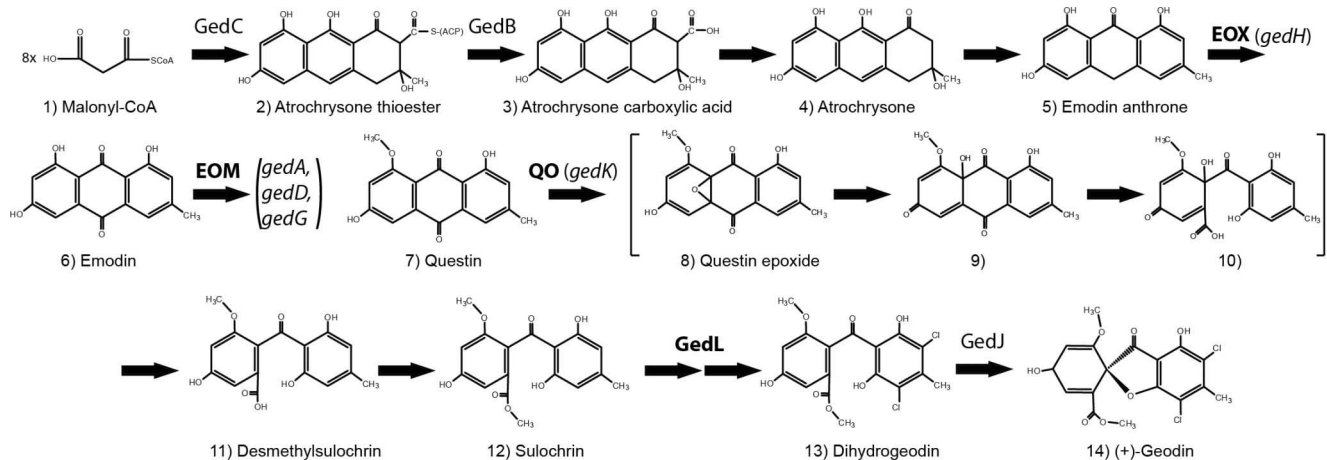


Figure 2. Proposed pathway for geodin production. The PKS (ACTS, [20]), thioesterase (ACTE, [20]) and dihydrogeodin oxidase previously linked to genes as well as the sulochrin halogenase identified in this study (highlighted in bold) are denoted by their *ged*-annotation. Enzymatic reactions for which the enzyme has been characterized but the gene not identified are marked in bold as EOX = emodin anthrone oxygenase, EOM = emodin-O-methyltransferase and QO = questin oxygenase. Reactions involving compounds 8-10 shown in brackets are inferred reactions proposed by Henry and Townsend based on a similar intra-molecular rearrangement in aflatoxin biosynthesis [26,27].

doi:10.1371/journal.pone.0072871.g002

Table 1. Summary of characterized and putative ORFs in the *A. terreus* geodin gene cluster.

BROAD annotation	Proposed annotation	Enzyme class	Validation	Reference
ATEG_08449	GedA	Putative O-methyl transferase*	BLAST, Hhpred	<i>This study</i>
ATEG_08450	GedB	β -lactamase type thioesterase	Heterologous expression, in vitro assays	[20]
ATEG_08451	GedC	Polyketide synthase	Deletion mutants, heterologous expression, in vitro assays	[19–21]
ATEG_08452	GedD	Putative O-methyl transferase*	BLAST, Hhpred	<i>This study</i>
ATEG_08453	GedR	Putative transcription factor	Deletion mutant, quantitative RT-PCR	<i>This study</i>
ATEG_08454	GedE	Putative glutathione-S-transferase	Annotation from BROAD, BLAST, Hhpred	<i>This study</i>
ATEG_08455	GedF	Putative oxidoreductase	BLAST, Hhpred	<i>This study</i>
ATEG_08456	GedG	Putative SAM-dependent-methyltransferase*	BLAST, Hhpred	<i>This study</i>
ATEG_08457-2	GedH	Putative emodin anthrone oxidase, similar to HypC	BLAST, 34% amino acid identity	<i>This study</i>
ATEG_08457-1	GedI	Putative mdpH homolog	BLAST, 46% amino acid identity	<i>This study</i>
ATEG_08458	GedJ	Dihydrogeodin oxidase	Enzymatic assays, protein sequencing	[18]
ATEG_08459	GedK	Putative Bayer Villiger-type oxidase	Enzymatic assays, Identity inferred in this study	[17]
ATEG_08460	GedL	Sulochrin halogenase	Deletion mutant, functional complementation	<i>This study</i>

All similarity percentages indicate identities at the amino acid level.

*One of these three putative ORFs is likely to encode the emodin O-methyltransferase described by Chen et al [16].

doi:10.1371/journal.pone.0072871.t001

annotated ATEG_08457.1) and ATEG_08457-2 (the new putative *HypC* homolog). Specifically, we suggest that ATEG_08457-2 and ATEG_08457-1 are positioned on *A. terreus* supercontig 12, base pairs 1307175–1307627 and 1308053–1308540, respectively.

Finally, we inspected the remaining ORFs in the region for activities relevant for production of geodin. Among these, one (ATEG_08454) was functionally annotated as a glutathione-S-transferase and two ORFs (ATEG_08455 and ATEG_08457-1) uncovered by the BLAST analysis displayed similarity to oxidoreductases and MdpH, respectively. The latter is a protein of unknown function required for emodin synthesis in *A. nidulans* [22]. To substantiate our predictions of the involvement of these putative genes, we conferred literature on similar biosynthetic pathways. The requirement for an oxidoreductase in geodin biosynthesis has previously been proposed by Henry and Townsend [26,27], while Simpson suggested the involvement of both a glutathione-S-transferase and an oxidoreductase in the biosynthesis of xanthenes in *A. nidulans* [28].

In addition to genes involved in the biosynthetic steps towards geodin, we noticed the presence of a gene, ATEG_08453, which encodes a putative transcription factor. The position of this gene within the putative geodin gene cluster suggests that it could regulate the activity of all genes in the cluster. In summary, our analysis suggests that the geodin gene cluster spans 25 kb and contains 13 putative ORFs (Figure 1A). Gene numbers, functional predictions and published data resulting from the entire analysis are presented in Table 1.

2.3 Strategy for transferring the geodin cluster from *A. terreus* into *A. nidulans*

Two vectors, containing 12 kb (Fragment 1) and 15 kb (Fragment 2) of the putative geodin gene cluster, respectively, were constructed by USER fusion by merging four individual PCR fragments in the first vector and seven PCR fragments in the second (Figure 1A). Assembling the geodin pathway from PCR fragments offers the possibility of introducing defined changes in the DNA sequence prior to integration via the many tools and methods for PCR based genetic engineering. In the present case,

we inserted the strong constitutive promoter, *PgpdA*, of *A. nidulans* upstream of ATEG_08453, which encodes the putative transcription factor described above, with the intention that this modification would activate transcription of all genes in the geodin cluster after its integration into *IS1*. Importantly, a fragment of the geodin cluster (2 kb) is included in both constructs to serve as the upstream targeting sequence for the second gene-targeting event as illustrated in Figure 1B.

2.4 Production of geodin in *A. nidulans*

Using the two vectors constructed above, the putative geodin gene cluster (*ged*) was successfully transferred to an *A. nidulans* reference strain as well as to a strain where the entire monodictyphenone gene cluster (*mdpA-LΔ*) had been deleted. Geodin production in the *mdpA-LΔ* strains would indicate that all genes in the geodin cluster are functionally expressed, while geodin formation in the reference strain could be mediated via metabolites produced in the monodictyphenone pathway. The ability of the recombinant strains to produce geodin on minimal medium was analyzed by UHPLC-HRMS. The presence of geodin in fungal extracts was identified by comparison of retention time, accurate mass spectra, and isotope ratio to an authentic geodin standard. These analyses demonstrated that both *ged*⁺ and *ged*⁺ *mdpA-LΔ* produced geodin (Figure 3A). Consequently, we suggest that the putative transcription factor, ATEG_08453, is renamed *gedR* and that the enzymes in the cluster (ATEG_08449–08452 + ATEG_08454–08460) are renamed *gedA-L*. In different experiments, we observed that the amount of geodin is reproducibly higher in the *ged*⁺ strain (40–70 μg/plate) as compared to the amount in the *ged*⁺ *mdpA-LΔ* strain (2–4 μg/plate), which does not produce emodin via the *mdp* cluster. We therefore speculate that natively produced emodin, or other intermediates towards emodin production, could be converted into geodin in the *ged*⁺-strains.

2.5 Genetic characterization of the geodin cluster in *A. nidulans*

One reason for transferring a gene cluster into a host with a well-developed genetic toolbox is the possibility for further

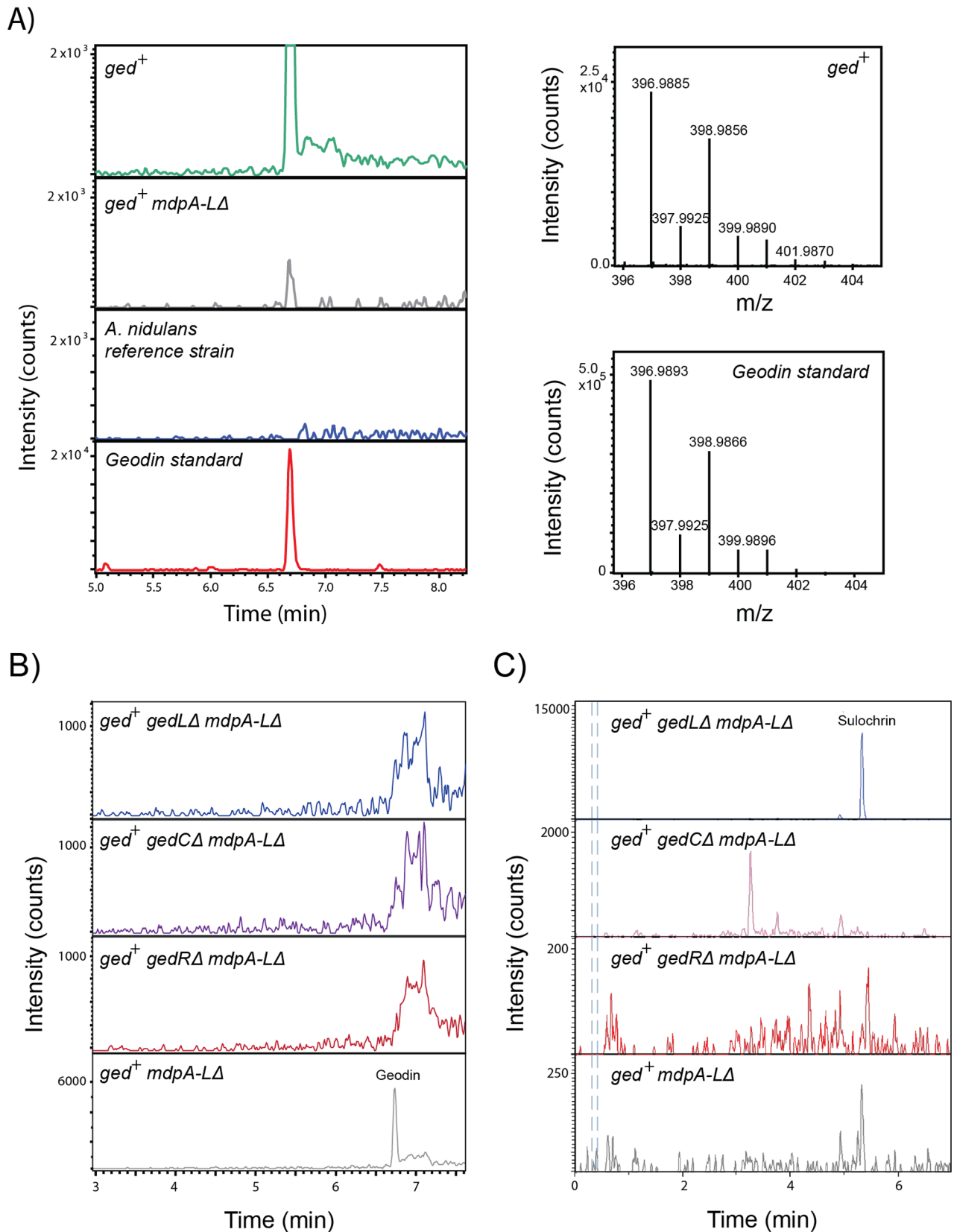


Figure 3. Production of geodin in *A. nidulans ged⁺* strains. **A)** Left panels depict extracted ion chromatograms (ESI-) of geodin m/z 396.9876 ± 0.005 amu from fungal extracts of *ged⁺*, *ged⁺ mdpA-LΔ* and reference strains (Bruker maXis system). An authentic geodin standard is included for comparison. The mass spectra of the putative geodin peak in *ged⁺* and the authentic geodin standard are depicted in panels to the right. **B)** and **C)** ESI- chromatograms of geodin m/z 396.9876 ± 0.005 amu (**B**) and sulochrin m/z 331.0812 ± 0.005 amu (**C**) extracted from *ged⁺ mdpA-LΔ* (grey), *ged⁺ mdpA-LΔ gedLΔ* (blue), *ged⁺ mdpA-LΔ gedCΔ* (purple) and *ged⁺ mdpA-LΔ gedRΔ* (red).
doi:10.1371/journal.pone.0072871.g003

characterization of the cluster. To demonstrate this possibility, we decided to investigate the functionality of three key genes in the cluster, *gedC*, *gedR* and *gedL* encoding the PKS, the putative regulator, and the putative halogenase, respectively. We focused our efforts on the *ged⁺ mdpA-Δ* strains, as they provide a genetic background with no risk of complementation by *mdp* enzymes. UHPLC-HRMS analysis of strains grown on minimal medium revealed that all three deletion strains were unable to synthesize geodin (Figure 3B), thereby confirming that geodin is indeed produced from the reconstituted cluster and that the corresponding proteins of all three genes were functional in *A. nidulans* and play a role in geodin biosynthesis. We note the presence of a co-eluting isobaric compound, seen as the broad peak (6.7–7.3 min) in Figure 3B. However, this compound is not geodin as it does not contain a chlorine isotopic pattern. In agreement with previous analyses [19–21], no intermediates of the proposed geodin pathway (Figure 2) accumulated in the *ged⁺ gedCΔ mdpA-Δ* strain, which is expected, as the PKS responsible for geodin formation is absent.

According to the proposed biosynthetic route for geodin production, the halogenase accepts sulochrin as substrate and adds two chlorine atoms to form dihydrogeodin (Figure 2). Consistent with the hypothesis that *gedL* encodes the sulochrin halogenase, sulochrin accumulated significantly in the *ged⁺ gedLΔ mdpA-Δ* strain (1.2 – 1.8 μg/plate), but was undetectable in the *gedR* or *gedC* deletion strains (Figure 3C). To confirm that this lack of halogenase activity was due to the *gedL* deletion we reintroduced the *gedL* ORF at another ectopic site, *IS3*, which is a site located on a chromosome different from the one harboring *IS1*, see Figure S2. Surprisingly, no production of geodin was observed in this strain (Figure S3A). This prompted us to perform a BLAST search of the GenBank database [23] using the amino acid sequence of the current ATEG_08460.1 gene model as query. Strikingly, the majority of the best hits were enzymes that contain additional 49 amino acid residues in their N-terminus, including a conserved MSIP/MSVP motif at the very N-terminal end, see Figure S4A. Interestingly, intron prediction based on Augustus [29] predicts an intron just upstream of the AUG proposed by the current gene model (ATEG_08460.1). Taking this into account and by using an ATG further upstream in the *gedL* gene, a very similar extension can be generated for GedL, see Figure S4B and C. We therefore inserted a larger fragment of the *gedL* locus that includes this new ATG as well as its native UTR sequence into *IS2* [4] in the *ged⁺ gedLΔ mdpA-Δ* strain. In this strain, geodin was produced in ample amounts (4.0 – 6.8 μg/plate) strongly suggesting that *gedL* indeed encodes the sulochrin halogenase. Interestingly, in this strain, targeted analysis of the UHPLC-HRMS data and comparison to an in-house metabolite database [30], revealed 0.04 – 0.06 μg/plate of sulochrin and trace amounts of monochlor-sulochrin indicating that chlorine is added in two discrete catalytic steps, see Figure S3B.

To investigate whether GedR regulates the genes of the geodin cluster in *A. nidulans*, we performed a gene specific mRNA transcript analysis by quantitative RT-PCR in the *ged⁺ mdpA-Δ* and *ged⁺ gedRΔ mdpA-Δ* strains for all genes in the geodin gene cluster where a putative homolog is present in the monodictyphenone cluster (Table 1). This analysis demonstrated that transcription of all seven selected genes was down regulated in the absence of GedR. Most prominently transcription from four of the genes (*gedF*, *G*, *H*, and *K*) was reduced to less than 10% of the level obtained in the *ged⁺ mdpA-Δ* strain (Figure 4). We note that the de novo annotated candidate gene for the emodin anthrone oxidase, *gedH* (ATEG_08457-2), is transcribed in both the *ged⁺ mdpA-Δ* and *ged⁺ mdp⁺* strains. In addition, its expression levels in the two

strains were different from those obtained for *gedI* (ATEG_08457-1). Together, these observations strongly indicate that *gedR* encodes a transcription factor, which activates the expression of the genes that are involved in geodin synthesis and that *gedH* is a genuine ORF.

Inspired by these results, we next tested whether GedR would activate the *gedR* promoter. To this end we inserted a *lacZ* reporter gene under the control of the native *gedR* promoter into *IS3*, in *ged⁺ mdpA-Δ* and *ged⁺ gedRΔ mdpA-Δ* strains. On MM medium containing 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-gal), colonies formed by the *Pgpda-lacZ* positive control strain were strongly blue, see Figure S5. The center of the colonies formed by *ged⁺ mdpA-Δ PgedR-lacZ* strain exhibited slightly blue color. However, this level of blue represents background as it did not differ from the amount and location of blue color produced by the negative control strain *ged⁺ mdpA-Δ*, see Figure S5. In agreement with this, a quantitative RT-PCR analysis showed that the *lacZ* mRNA level was only modestly increased (1.5 fold) in the *ged⁺ mdpA-Δ PgedR-lacZ* strain as compared to a *ged⁺ gedRΔ mdpA-Δ PgedR-lacZ* strain, but this difference was not statistically significant ($p = 0.08$). Thus, GedR is not sufficient to induce expression from *gedR* in *A. nidulans*.

The fact that geodin production was significantly higher in the *ged⁺* than in the *ged⁺ mdpA-Δ* strains prompted us to investigate whether GedR could also activate transcription of the *mdp* cluster. Specifically, we compared transcription from *mdpG*, encoding the monodictyphenone PKS, in the *ged⁺* and the reference strains. In agreement with our hypothesis the *mdpG* transcript was easily detectable in the *ged⁺* strain, but undetectable in the reference, see Figure S6.

2.6 Conservation of gene clusters resembling the geodin cluster in other fungal species

Finally, we speculated whether gene clusters of a similar organization could be found in other sequenced fungal species as emodin is well-known to serve as precursor to a wide range of natural products [31–34]. Comparison of the geodin gene cluster to all *Aspergillus* genomes available at the *Aspergillus* Comparative Sequencing Project database (Broad Institute of Harvard and MIT, <http://www.broadinstitute.org/>) revealed the presence of putative gene clusters in *A. fumigatus* and *A. fischerianus* containing putative homologs of 12 of the 13 annotated ORFs in the geodin

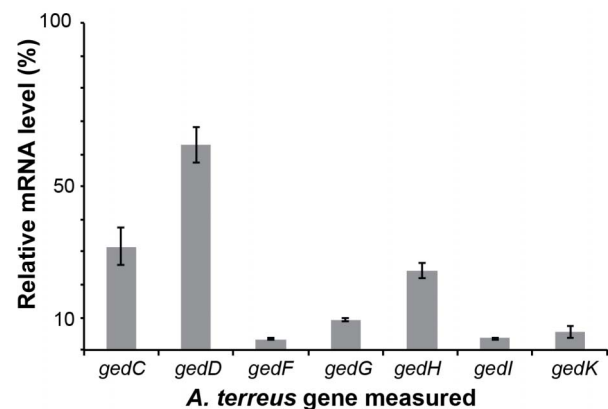


Figure 4. The *A. terreus* transcription factor GedR is important for gene expression in the geodin gene cluster in *A. nidulans*. Transcription levels of selected *ged*-genes in *ged⁺ mdpA-Δ*, *gedRΔ mdpA-Δ* strains relative to the corresponding levels in *ged⁺ mdpA-Δ* strains. doi:10.1371/journal.pone.0072871.g004

cluster (the halogenase, *gedL*, is absent). The internal organization of the putative clusters in *A. fumigatus* (Afu4g14450-14580) and *A. fischerianus* (101790-101920) were identical to the geodin cluster with the exception of an inversion affecting the five ORFs *gedG*-*gedK*. Moreover, the amino acid identities between biosynthetic enzymes in *A. terreus* and *A. fumigatus*/*A. fischerianus* were in average 58% and 60%, respectively. The conservancy across these three species further substantiates our delineation of the geodin cluster and hints that the putative clusters in *A. fumigatus* and *A. fischerianus* may encode the biosynthesis for a similar compound. Both species are known to produce trypacidin [35], which differs from geodin only by the absence of chlorines and the presence of an additional methyl group [36]. In agreement with the structural differences between geodin and trypacidin, the putative *A. fumigatus* and *A. fischerianus* gene clusters contain one additional putative methyltransferase, but lack the putative halogenase. Thus, the two putative clusters are candidates for trypacidin gene clusters.

Materials and Methods

Strains and media

Escherichia coli strain DH5 α was used to propagate all plasmids. Genomic DNA from the geodin producing *A. terreus* IBT15722 strain was used as template for PCR amplification of the geodin cluster. *A. nidulans* strains are shown in Table S1. *A. nidulans* strains were grown on solid glucose minimal medium (MM) prepared as described by Cove [37], but with 1 % glucose, 10 mM NaNO₃ and 2 % agar. MM was supplemented with 10 mM uridine (Uri), 10 mM uracil (Ura), and/or 4 mM L-arginine (Arg) when required. Solid plates containing 5-fluoroorotic acid (5-FOA) were made as MM+Uri+Ura medium supplemented with filter sterilized 5-FOA (Sigma-Aldrich) to a final concentration of 1.3 mg/ml.

Vector construction

All vectors were made by USER cloning and USER fusion [7,11]. All PCR products were amplified in 35 cycles using proof-reading PfuX7 polymerase [38]. Next USER fusions of vector and inserts were performed as previously described [12]. Reactions were incubated for 20 min at 37 °C, followed by 20 min at 25 °C before transformation into *E. coli*.

The pU2111-3 vector was constructed by USER fusion of 5 PCR amplified fragments: 1) vector backbone for propagation in *E. coli* (amplified with primers DH110/DH111), 2) US (upstream) targeting sequence for insertion in *IS3* (DH112/DH113), 3) PgpA-S::UEC::TrpC (DH114/DH115), 4) *A. fumigatus pyrG* (marker) (DH116/DH117) and 5) DS (downstream) targeting sequence for insertion in *IS3* (DH118-DH119). UEC: uracil excision cassette. Template for fragments 1 and 3: pU1111, for fragments 2 and 5: *A. nidulans* genomic DNA, and for fragment 4: pDEL2 [13]. p2110-3-*lacZ* was constructed by combining an AsiSI and Nb.BtsI pU2111-3 vector fragment with a PCR product containing the *E. coli lacZ* gene (amplified from pU2110-1-*lacZ* using motn136/motn137 [7] as primers) by USER cloning. The plasmid p2010-3-PgedR-*lacZ* was constructed by USER fusion of 5 PCR amplified fragments: 1) *gedR* promoter sequence (DH120/DH121, template: NID677 genomic DNA with), 2) *lacZ*::*TrpC*::*AFpyrG* (motn136/DH122, template: p2111-3), 3) vector backbone for propagation in *E. coli* (DH110/DH111 template: pU1111), 4) DS targeting sequence for insertion in *IS3* (DH123/DH119 template: pU2111-3) and 5) US targeting sequence for insertion in *IS3* (DH112/DH113). The vectors pU2110-3-ATEG_08460.1 and pU2110-2-*gedL* were constructed by combining the PCR

fragments generated with the primers JBN K35/K36 and JBN W77/W78 with pU2111-3 and pU2111-2 [4] vector fragments, respectively, by USER fusion. Prior to USER fusion both plasmids were digested with AsiSI and Nb.BtsI to create the vector fragments.

The two integration vectors containing the geodin gene cluster, pU1110-1-*ged1* (containing Fragment 1) and pU2000-*ged2* (containing Fragment 2), were made as follows. Primers for generating all PCR fragments for Fragment 1 and 2 assemblies are shown in Table S2. pU1110-1-*ged1* was constructed by combining all relevant fragments for Fragment 1 assembly into an AsiSI and Nb.BtsI pU1111-1 vector fragment by USER fusion. The vector fragment of pU2000-*ged2* is based on two PCR fragments generated by using primers 77/422 and 421/70 as well as pJ204 [7] and pU2111-1, respectively, as templates. These two PCR products and all relevant fragments for Fragment 2 assembly were then combined by USER fusion. The inserts in the two integration vectors were fully sequenced (StarSEQ, Germany).

Construction of *A. nidulans* strains

Protoplastation and gene-targeting procedures were performed as described by Nielsen et al [14] using either *argB* or *AFpyrG* as marker. All strains were verified by PCR analysis using spores as the source of DNA. Prior to PCR, the samples were incubated for 25 min at 98°C to liberate genomic DNA. This treatment was followed by touch-down PCR programs with annealing temperatures ranging from 64–56°C. Reactions were carried out in 35 cycles using 40 μ L volume with less than 1000 spore (one light stab in the colony with a pipette tip).

The *ged*⁺ (NID677) and *ged*⁺ *mdpA-LA* (NID695) strains were obtained by transformation of the relevant gene targeting substrates into NID74 [39] or NID356, respectively. The gene targeting substrates containing Fragment 1 and Fragment 2 were liberated from pU1111 (NotI) and pU2052 (SwaI), respectively, and gel purified (GFXTM, GE Healthcare) prior to transformation. The resulting strains, NID677 and NID695, were subjected to counter selection on 5-FOA, generating NID802 and NID823, in order to recycle the *AFpyrG* marker, hence, allowing the use of this marker for subsequent gene deletions of *gedC*, *gedL* and *gedR*. These deletions were made as described in [14] using the primers listed in Table S2.

The strains NID1291 and NID1297, expressing *lacZ* under the control of ATEG_08453 promoter was made by transforming a gene targeting substrate liberated from pU2010-3-PgedR-*lacZ* by digestion with SwaI into NID823 and NID925, respectively. A control strain expressing *lacZ* under the constitutive promoter *PgpA* NID1278, was constructed by integrating the gene-targeting substrate liberated from pU2110-3-*lacZ* by SwaI. Both gene-targeting substrates integrate into *IS3* located between genes AN4770 and AN4769 on chromosome III, at position 1047840-1051735, by homologous recombination (See Figure S2).

The halogenase complementation strains, ATEG_08460.1 and *gedL*, were made in the NID1279 background, a pop-out recombinant strain from NID843. Digestion of vector pU2110-3-ATEG_08460.1 and pU2110-2-*gedL* liberated gene-targeting constructs for transformation and integration in *IS3* and *IS2*, respectively.

Chemical characterization of *A. nidulans* strains

All strains were grown as three point stab inoculations for 7 days at 37°C in the dark on solid MM-media. Extraction and analysis of

metabolites were performed by 2 methods: i) The agar plug extraction method described by Smedsgaard [40], using a total of 1 cm² of a colony, followed by analysis using reversed phase separation UHPLC-UV/VIS-HRMS on a maXis G3 quadrupole time of flight (qTOF) mass spectrometer (Bruker Daltonics, Bremen, Germany) connected to an Ultimate 3000 UHPLC system (Dionex, Sunnyvale, CA), and equipped with a 10 cm Kinetex C₁₈ column (Phenomenex Torrance, Ca, USA) running a 10–100% acetonitrile gradient system in 10 min at 40°C; ii) more concentrated samples were made by extracting metabolites from a total of 15 cm² of a colony using 12 ml solvent (ethyl acetate-dichloromethane-methanol-formic acid 60:30:15:1 v/v/v) in a 16-ml vial. The extract was evaporated to dryness with N₂ flow and redissolved in 0.5 ml methanol and analyzed by reversed phase separation on Agilent 1290 UHPLC coupled to an Agilent 6550 qTOF (Santa Clara, CA, USA) equipped with an electrospray source, and equipped with a 25 cm Agilent Poroshell phenyl hexyl running a 10–100% acetonitrile gradient system in 15 min 60°C. Both MS instruments were mainly operated in ESI⁺ as geodin and related compounds ionizes best here in this mode [30]. Identification and quantification of geodin and sulochrin (BioAustralis, Smithfield, NSW, Australia) were based on comparison of peak area, retention time, accurate mass (± 1.5 ppm), isotope pattern and adduct pattern to quantitative authentic standards. Non quantitative standards representing 5-O-methylsulochrin; sulochrin-2'-methylether; isosulochrin; 3-O-demethylsulochrin; tryptacidin; and emodin were also included in the analyses. Other intermediates were identified by comparison to an internal reference standard database (~1500 compounds) [30]. For the identification of geodin in NID695, high resolution MS (50 000 FWHM) and mass accuracy (< 1.5 ppm) of the maXis G3 was needed to exclude a non chlorine containing co-eluting isobaric compound, seen as the broad peak in Figure 3 (6.7–7.3 min) that impaired the identification of geodin. In the following strains geodin was further verified by better chromatographic separation on a 25 cm phenyl-hexyl column on the Agilent UHPLC-qTOF.

RNA isolation and quantitative RT-PCR

RNA isolation from the *A. nidulans* strains and subsequent quantitative RT-PCR reactions were done as previously described in [7] except that biomass for RNA isolation was prepared with a Tissue-Lyser LT (Qiagen) by treating samples for 1 min at 45 MHz. The *A. nidulans* histone 3 encoding gene, *hhtA* (AN0733) was used as an internal standard for normalization of expression levels. All primers used for quantitative RT-PCR are shown in Table S2.

Bioinformatic analysis of *gedL* (ATEG_08460.1)

Alignment illustrations and sequence were made in CLC Main Workbench 6.8.4.; Alignment parameters Gap open costs = 10, Gap extension cost = 1. All sequences in the alignments can be retrieved from Genbank, by the following references for the putative halogenases: *Aspergillus oryzae* ref|XP_001818590.1|, *Aspergillus terreus* ref|XP_001217599.1| and the *A. terreus* genome sequence resource at the Broad Institute, *Bipolaris sorokiniana* gb|EMD66881.1|, *Chaetomium chiversii* (*RadH*) gb|ACM42402.1|, *Chaetomium globosum* ref|XP_001227515.1|, *Pochonia chlamydosporia* (Rdc2) gb|ADM86580.1| [39] and *Talaromyces stipitatus* ref|XP_002486044.1|. Intron prediction of the *A. terreus* ATEG_08460.1 locus was done based on the Augustus gene prediction resource [29].

Concluding Remarks

We have described the complete and targeted transfer of all 13 genes of the geodin gene cluster from *A. terreus* to *A. nidulans* through a sequential integration approach enabling *A. nidulans* to synthesize geodin. In principle, this strategy can be used to reconstitute gene clusters of any size as the sequential integrations are based on marker recycling. In addition, defined promoters can easily be introduced in front of relevant genes in the cluster of interest. Importantly, we demonstrate that the cluster can be genetically dissected for clarification of its biochemical potential. We therefore envision that our method will significantly speed up the uncovering of biochemical pathways in fungi where the genome has been sequenced.

Supporting Information

Figure S1 Identification of putative HypC homolog encoded by *gedH* (ATEG_08457-2) in the *A. terreus* geodin gene cluster. Pairwise alignment of putative emodin anthrone oxidase, GedH (ATEG_08457-2), from *A. terreus* and norsolinic anthrone oxidase, HypC, from *A. flavus*. The conserved DUF-1772 domain and putative catalytic regions proposed by Ehrlich et al [25] are highlighted in green and blue, respectively. (TIF)

Figure S2 Schematic overview of the integration of a gene-expression cassette into the integration site, *IS3*, by homologous recombination. *IS3* is located between genes AN4770 and AN4769 on chromosome III. The cassette consists of six parts: upstream targeting sequence (US), promoter (P, in this case 0.5 kb *PgpdA*), your favorite gene (YFG), terminator (T, *TtpC*), marker (in this case *AFpyrG* flanked by direct), and the downstream targeting sequence (DS). The orientations of the genes AN4770 and AN4769 are indicated by green arrows. The sizes of US, DS and the intergenic region are 1984 bp, 1911 bp, and 3007 bp, respectively. (TIF)

Figure S3 Complementation of halogenase deficiency. **A)** Left panel: detection of sulochrin (-ESI, EIC(*m/z* 331.0812)); right panel: detection of geodin (-ESI, EIC(*m/z* 396.9876)). Strains for halogenase analysis, from top to bottom: NID843 (*gedLA*); NID1280 (*gedLA*, *IS3::PgpdA*-ATEG_08460.1); and NID1306 (*gedLA*, *IS2::gedL*). **B)** Ratio of geodin, monochlor-sulochrin, and sulochrin in the NID1306 strain, including ESI⁺-MS spectrum of monochlor-sulochrin showing the isotopic pattern and the mass deviations relative to the theoretical masses. Reference standards of geodin and sulochrin were included in all runs (data not shown). (TIF)

Figure S4 Identification of the likely start codon of *gedL*. **A)** Alignment of the top hits in a BLAST search for ATEG_08460.1 homologs shows that they contain a very conserved 48 amino acid residue addition in the N-terminus. Amongst the homologs, Rdc2, has been characterized as a halogenase by [39] MLAS is the predicted N-terminus of ATEG_08460.1. Drawing is not to scale. **B)** The position of putative exons and intron in the 5'-end of *gedL* as predicted by the Augustus software [29]. The predicted protein sequence encoded by exon 1 and by the first section of exon 2 is indicated. **C)** Full alignment of the halogenase homologs and GedL based on the GedL sequence derived from the new start codon. (TIF)

Figure S5 Expression of *lacZ* under the control of the *gedR* promoter (*PgedR*). Left panel: the positions of the strains

on the plate are shown in the right panel. NID823 (*ged*⁺ *mdpA-LΔ*) is the reference strain without the *lacZ* gene. NID1278 is a control strain containing the *P_{gpdA}-lacZ* construct in *IS3*. The NID1291 (*ged*⁺ *mdpA-LΔ* *P_{gedR}-lacZ*) strain carries *P_{gedR}-lacZ* in *IS3*. The strains were stabbed on MM containing X-gal and incubated three days at 37 °C in the dark before photography. (TIF)

Figure S6 Constitutive expression of *gedR* induces transcription of the *A. nidulans* gene *mdpG*. *mdpG* mRNA levels in reference (NID1) and in the *ged*⁺ strain (NID677) were evaluated by quantitative RT-PCR. For each strain, RNA was extracted as described in Materials and Method and the RNA samples analyzed in triplicate by quantitative RT-PCR. The samples were loaded and analyzed by 1% agarose gel-electrophoresis as indicated in the figure. (TIF)

Table S1 Strain genotypes. *= For reference see Nielsen et al ([13]). (XLSX)

References

- Wilson RA, Talbot NJ (2009) Fungal physiology - a future perspective. *Microbiology-Sgm* 155: 3810–3815.
- Hoffmeister D, Keller NP (2007) Natural products of filamentous fungi: enzymes, genes, and their regulation. *Natural Product Reports* 24: 393–416.
- Keller NP, Hohn TM (1997) Metabolic pathway gene clusters in filamentous fungi. *Fungal Genetics and Biology* 21: 17–29.
- Hansen BG, Mnich E, Nielsen KF, Nielsen JB, Nielsen MT, et al (2012) Involvement of a natural fusion of a cytochrome P450 and a hydrolase in mycophenolic acid biosynthesis. *Applied and Environmental Microbiology* 78: 4908–4913.
- Hengghan MN, Yakasai AA, Halo LM, Song ZS, Bailey AM, et al (2010) First heterologous reconstruction of a complete functional fungal biosynthetic multigene cluster. *Chembiochem* 11: 1508–1512.
- Itoh T, Tokunaga K, Matsuda Y, Fujii I, Abe I, et al (2010) Reconstitution of a fungal meroterpenoid biosynthesis reveals the involvement of a novel family of terpene cyclases. *Nature Chemistry* 2: 858–864.
- Hansen BG, Salomonsen B, Nielsen MT, Nielsen JB, Hansen NB, et al (2011) Versatile enzyme expression and characterization system for *Aspergillus nidulans*, with the *Penicillium brevicompactum* polyketide synthase gene from the mycophenolic acid gene cluster as a test case. *Applied and Environmental Microbiology* 77: 3044–3051.
- Wingler LM, Cornish VW (2011) Reiterative recombination for the in vivo assembly of libraries of multigene pathways. *Proceedings of the National Academy of Sciences of the United States of America* 108: 15135–15140.
- Bergmann S, Schumann J, Scherlach K, Lange C, Brakhage AA, et al (2007) Genomics-driven discovery of PKS-NRPS hybrid metabolites from *Aspergillus nidulans*. *Nature Chemical Biology* 3: 213–217.
- Sakai K, Kinoshita H, Shimizu T, Nihira T (2008) Construction of a citrinin gene cluster expression system in heterologous *Aspergillus oryzae*. *Journal of Bioscience and Bioengineering* 106: 466–472.
- Nour-Eldin HH, Hansen BG, Norholm MH, Jensen JK, Halkier BA (2006) Advancing uracil-excision based cloning towards an ideal technique for cloning PCR fragments. *Nucleic Acids Research* 34: e122.
- Geu-Flores F, Nour-Eldin HH, Nielsen MT, Halkier BA (2007) USER fusion: a rapid and efficient method for simultaneous fusion and cloning of multiple PCR products. *Nucleic Acids Research* 35: e55.
- Nielsen JB, Nielsen ML, Mortensen UH (2008) Transient disruption of non-homologous end-joining facilitates targeted genome manipulations in the filamentous fungus *Aspergillus nidulans*. *Fungal Genetics and Biology* 45: 165–170.
- Nielsen ML, Albertsen L, Lettier G, Nielsen JB, Mortensen UH (2006) Efficient PCR-based gene targeting with a recyclable marker for *Aspergillus nidulans*. *Fungal Genetics and Biology* 43: 54–64.
- Chen ZG, Fujii I, Ebizuka Y, Sankawa U (1995) Purification and characterization of emodinanthrone oxygenase from *Aspergillus terreus*. *Phytochemistry* 38: 299–305.
- Chen ZG, Fujii I, Ebizuka Y, Sankawa U (1992) Emodin O-methyltransferase from *Aspergillus terreus*. *Archives of Microbiology* 158: 29–34.
- Fujii I, Ebizuka Y, Sankawa U (1988) A novel anthraquinone ring cleavage enzyme from *Aspergillus terreus*. *Journal of Biochemistry* 103: 878–883.
- Huang KX, Fujii I, Ebizuka Y, Gomi K, Sankawa U (1995) Molecular cloning and heterologous expression of the gene encoding dihydrogeodin oxidase, a multicopper blue enzyme from *Aspergillus terreus*. *Journal of Biological Chemistry* 270: 21495–21502.
- Askenazi M, Driggers EM, Holtzman DA, Norman TC, Iverson S, et al (2003) Integrating transcriptional and metabolite profiles to direct the engineering of lovastatin-producing fungal strains. *Nature Biotechnology* 21: 150–156.
- Awakawa T, Yokota K, Funa N, Doi F, Mori N, et al (2009) Physically discrete beta-lactamase-type thioesterase catalyzes product release in atrochrysone synthesis by iterative type I polyketide synthase. *Chemistry & Biology* 16: 613–623.
- Couch RD, Gaucher GM (2004) Rational elimination of *Aspergillus terreus* sulochrin production. *Journal of Biotechnology* 108: 171–178.
- Chiang YM, Szweczyk E, Davidson AD, Entwistle R, Keller NP, et al (2010) Characterization of the *Aspergillus nidulans* monodictyphenone gene cluster. *Applied and Environmental Microbiology* 76: 2067–2074.
- Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic Local Alignment Search Tool. *Journal of Molecular Biology* 215: 403–410.
- Soding J, Biegert A, Lupas AN (2005) The HHpred interactive server for protein homology detection and structure prediction. *Nucleic Acids Research* 33: W244–W248.
- Ehrlich KC, Li P, Scharfenstein L, Chang PK (2010) HypC, the anthrone oxidase involved in aflatoxin biosynthesis. *Applied and Environmental Microbiology* 76: 3374–3377.
- Henry KM, Townsend CA (2005) Synthesis and fate of o-carboxybenzophenones in the biosynthesis of aflatoxin. *Journal of the American Chemical Society* 127: 3300–3309.
- Henry KM, Townsend CA (2005) Ordering the reductive and cytochrome P450 oxidative steps in demethylsterigmatocystin formation yields general insights into the biosynthesis of aflatoxin and related fungal metabolites. *Journal of the American Chemical Society* 127: 3724–3733.
- Simpson TJ (2012) Genetic and biosynthetic studies of the fungal prenylated xanthone shamixanthone and related metabolites in *Aspergillus* spp. revisited. *Chembiochem* 13: 1680–1688. 10.1002/cbic.201200014.
- Stanke M, Morgenstern B (2005) AUGUSTUS: a web server for gene prediction in eukaryotes that allows user-defined constraints. *Nucleic Acids Research* 33: W465–W467.
- Nielsen KF, Mansson M, Rank C, Frisvad JC, Larsen TO (2011) Dereplication of microbial natural products by LC-DAD-TOFMS. *Journal of Natural Products* 74: 2338–2348.
- Gatenbeck S, Malmström L (1969) On the biosynthesis of sulochrin. *Acta Chemica Scandinavica* 23: 3493–3497.
- Steglich W, Arnold R, Losel W, Reininger W (1972) Biosynthesis of anthraquinone pigments in *Dermocybe*. *J Chem Soc, Chem Commun* 102–103.
- Franck B, Huper F (1966) A new principle of biosynthesis involving secoanthraquinones. *Angew Chem Int Ed Engl* 5: 728–729. 10.1002/anie.196607282.
- Sankawa U, Ebizuka Y, Shibata S (1973) Biosynthetic incorporation of emodin and emodinanthrone into the anthraquinonoids of *Penicillium brunneum* and *P. islandicum*. *Tetrahedron Letters* 14: 2125–2128.
- Larsen TO, Smedsgaard J, Nielsen KF, Hansen MAE, Samson RA, et al (2007) Production of mycotoxins by *Aspergillus lentulus* and other medically important and closely related species in section Fumigati. *Med Mycol* 45: 225–232.
- Afzal M, Davies JS, Hassall CH (1969) The biosynthesis of phenols. Part XIX. Synthesis of the polyhydroxybenzophenone derivative, sulochrin, of the gris-

- 2',5'-diene-3,4'-dione, tryptacin, and of related compounds. *J Chem Soc C* 1721–1727.
37. Cove DJ (1966) Induction and Repression of Nitrate Reductase in Fungus *Aspergillus Nidulans*. *Biochimica et Biophysica Acta* 113: 51–&.
 38. Norholm MHH (2010) A mutant Pfu DNA polymerase designed for advanced uracil-excision DNA engineering. *Bmc Biotechnology* 10: 10.1186/1472-6750-10-21.
 39. Nielsen MT, Klejnstrup ML, Rohlfis M, Anyaogu DC, Nielsen JB, et al (2013) *Aspergillus nidulans* synthesizes insect juvenile hormones upon expression of a heterologous regulatory protein and in response to grazing by *Drosophila melanogaster* larvae. *Plos One* 10.1317/journal.pone.0073369.
 40. Smedsgaard J (1997) Micro-scale extraction procedure for standardized screening of fungal metabolite production in cultures. *Journal of Chromatography A* 760: 264–270.

Chapter 12

Over-expression of Transcription Factor in *Aspergillus aculeatus* Activates Biosynthetic Pathway and Reveals Novel Compounds Acurin A and B

This chapter contains a manuscript in preparation, "Over-expression of Transcription Factor in *Aspergillus aculeatus* Activates Biosynthetic Pathway and Reveals Novel Compounds Acurin A and B", of which I am joint first author. The chapter describes the investigation of a putative 6-MSA synthase in *A. aculeatus*, and the products to which 6-MSA is a precursor. I conceived the idea for this study based on an analysis of all synthases identified from the genome sequence of *A. aculeatus* ATCC16872, which is publicly available from the DOE Joint Genome Institute¹. Using a combination of BLAST analyses with known PKS as query sequences, and keyword searches for specific enzymatic activities, I identified a total of 57 synthases: 24 PKSs, 19 NRPSs, nine terpenoid synthases and five hybrids. Details of this study is presented in Appendix H. The PKSs comprised five NR-PKSs, 17 partly or fully reducing PKSs, three non-ordinary PKSs, and interestingly: a putative 6-MSA synthase.

Heterologous expression of the gene verified that it encoded a 6-MSAS synthase. The AMA1-based expression platform (described in Section 3.1.4) was applied for over-expression of the PKS gene and a TF located upstream of the 6-MSA synthase, resulting in production of two novel compounds. A total of five compounds were linked to this gene cluster, and feeding studies with fully ¹³C labeled 6-MSA verified that 6-MSA was incorporated into three of the compounds.

¹<http://genome.jgi-psf.org/Aspac1/Aspac1.home.html>

At present time, we have not elucidated the structures of the three compounds that incorporate 6-MSA. Thus, we aim to solve the structures before submission, enabling a discussion about biosynthetic routes and origins and possible relationships to the novel compounds acurin A and B.

My contribution to this study was conception of the idea, experimental design of genetic work and feeding studies, supervision of and execution of genetic work, execution of feeding studies, and analysis of data from feeding studies. The manuscript was written in collaboration with L.M.P. Supporting information is included in Appendix G.

Over-expression of TF in *A. aculeatus* reveals novel compounds acurin A and B

Over-expression of Transcription Factor in *Aspergillus aculeatus* Activates Biosynthetic Pathway and Reveals Novel Compounds Acurin A and B

Lene M. Petersen^{†,1}, Dorte K. Holm^{†,2}, Charlotte H. Gotfredsen³, Uffe H. Mortensen², Thomas O. Larsen^{1,*}

Affiliation:

¹Chemodiversity group, Center for Microbial Biotechnology, Department of Systems Biology, Soltofts Plads, Building 221, Technical University of Denmark, 2800 Kgs. Lyngby, Denmark.

²Eukaryotic Molecular Cell Biology group, Center for Microbial Biotechnology, Department of Systems Biology, Soltofts Plads, Building 223, Technical University of Denmark, 2800 Kgs. Lyngby, Denmark.

³Department of Chemistry, Kemitorvet, Building 201, Technical University of Denmark, 2800 Kgs. Lyngby, Denmark.

[†]These authors have contributed equally to this work.

Over-expression of TF in *A. aculeatus* reveals novel compounds acurin A and B

***Corresponding author:**

Thomas O. Larsen, Technical University of Denmark, Department of Systems Biology

Center for Microbial Biotechnology, Building 221

2800 Kgs. Lyngby, Denmark

Email: tol@bio.dtu.dk

Phone: +45 45252632

Fax: +45 45884922

ABSTRACT

Aspergillus aculeatus is a filamentous fungus, belonging to the *Aspergillus* clade *Nigri*, which is an industrial workhorse for enzyme production. Only a few secondary metabolites have been reported for this fungus, however, its genetic potential for production of secondary metabolites is vast. In this study we have identified a 6-methylsalicylic acid (6-MSA) synthase from *A. aculeatus*, and verified its functionality by episomal expression in *A. aculeatus* and heterologous expression in *A. nidulans*. Furthermore, we have over-expressed a transcription factor, embedded in the gene cluster, and found that it activated a biosynthetic cluster and produced two novel compounds, which we have named acurin A and B. Feeding studies with fully ¹³C-labeled 6-MSA dismisses that 6-MSA is incorporated into acurin A and B, however, 6-MSA is incorporated into three other novel, and perhaps related, compounds.

INTRODUCTION

Aspergillus aculeatus is a filamentous fungus belonging to the group of black Aspergilli. *A. aculeatus* was recently sequenced by the DOE Joint Genome Institute (JGI), but no genes have yet been linked to production of metabolites. *A. aculeatus* is a known producer of several important secondary metabolites including okaramines [1], the toxic secalonic acids [2] and recently we have reported production of the antifungal compounds calbistrins as well as several novel compounds with yet unknown activities[3]. The latter include acu-morpholine and a group of related compounds of mixed biosynthetic origin, the aculenes.

Analysis of the published genome sequence revealed that *A. aculeatus* contains a gene with high similarity to 6-MSA synthases, however, the gene cluster containing the putative 6-MSA synthase does not resemble any reported gene clusters containing a 6-MSAS synthase e.g. the *yde* cluster in *A. niger* [4] and the *pat* cluster in *A. terreus* [5]. No 6-MSA or obvious 6-MSA-derived compounds are detected in wildtype extracts of *A. aculeatus*, so it is possible that 6-MSA is highly modified into the final metabolite as it is the case for both yanuthone D [4] and patulin [5].

As no genetic techniques exist for *A. aculeatus*, we have used an AMA1-based expression platform for gene expression. AMA1-containing plasmids allow for autonomous replication in many different Aspergilli [6,7] why this is a quick and efficient tool for expressing genes in organisms that do not have available genetic techniques. Furthermore, the transformation efficiency with autonomous plasmids is much higher than integrating linear substrates [6] why it is also an efficient mean for gene expression in organisms that protoplast or transform very poorly.

In this study we have shown that the putative 6-MSA synthase is in fact a 6-MSA synthase, by heterologous expression of the gene in the model fungus *Aspergillus nidulans*. Furthermore, we have used the AMA1-based expression platform in *A. aculeatus* for expression of a transcription factor embedded in the 6-MSA gene cluster. This approach activated a biosynthetic cluster, which is responsible for production of five novel compounds, including acurin A and B, which we report for the first time in this study.

RESULTS AND DISCUSSION

Heterologous expression of putative 6-MSA synthase verifies production of 6-MSA. A BLAST analysis of the 1904397 gene on all genomes in the Aspergillus Genome Database (AspGD) revealed a high similarity to fungal 6-methylsalicylic acid (6-MSA) synthases. To verify that the gene encodes a 6-MSA synthase the gene was integrated into the defined and characterized locus *IS1* of *A. nidulans*, which is a site that was previously used for successful expression of PKS genes [4, 8]. This approach is useful in this study, as *A. nidulans* is known to not produce 6-MSA. The metabolite profile of the strain was analyzed by UHPLC-DAD-TOFMS, and a new compound was produced compared to the reference (Figure 1). Dereplication of the compound and comparison to an authentic standard of 6-MSA verified that 6-MSA (**1**) was indeed produced, confirming that 1904397 encodes a 6-MSA synthase.

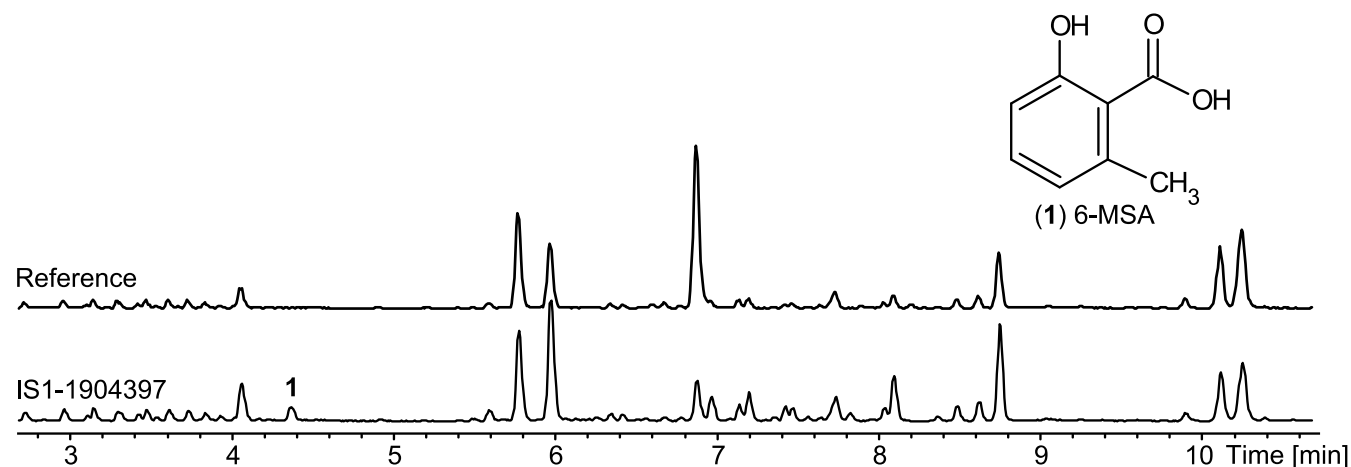


Figure 1. BPC of the *A. nidulans* strain expressing the putative 6-MSA synthase 1904397 compared to an *A. nidulans* reference.

Expression of 6-MSA synthase in *A. aculeatus* results in accumulation of 6-MSA. 6-MSA was not detected in wildtype extracts of *A. aculeatus*, why we hypothesized that over-expression of the 6-MSA synthase could result in production of a new compound, or in higher amounts of a known compound, thus identifying the end product of the biosynthetic pathway. Accordingly, the 6-MSA synthase was expressed in *A. aculeatus* from the plasmid pDHX3, carrying the *hph* selection marker. Correct transformants were cultivated on solid minimal medium with hygromycin B and subjected to metabolite profile analysis by UHPLC-DAD-TOFMS. Data analysis verified that the strain produced 6-MSA (Figure 2). However, there was no significant change in the rest of the metabolite profile (Figure 2). The fact that the *A. aculeatus* pDHX3-6-MSAS strain produced 6-MSA rather than a biosynthetic end product, indicated that any tailoring enzymes in the biosynthetic pathway are not expressed. Thus, the entire 6-MSA cluster may be silent under the applied laboratory conditions.

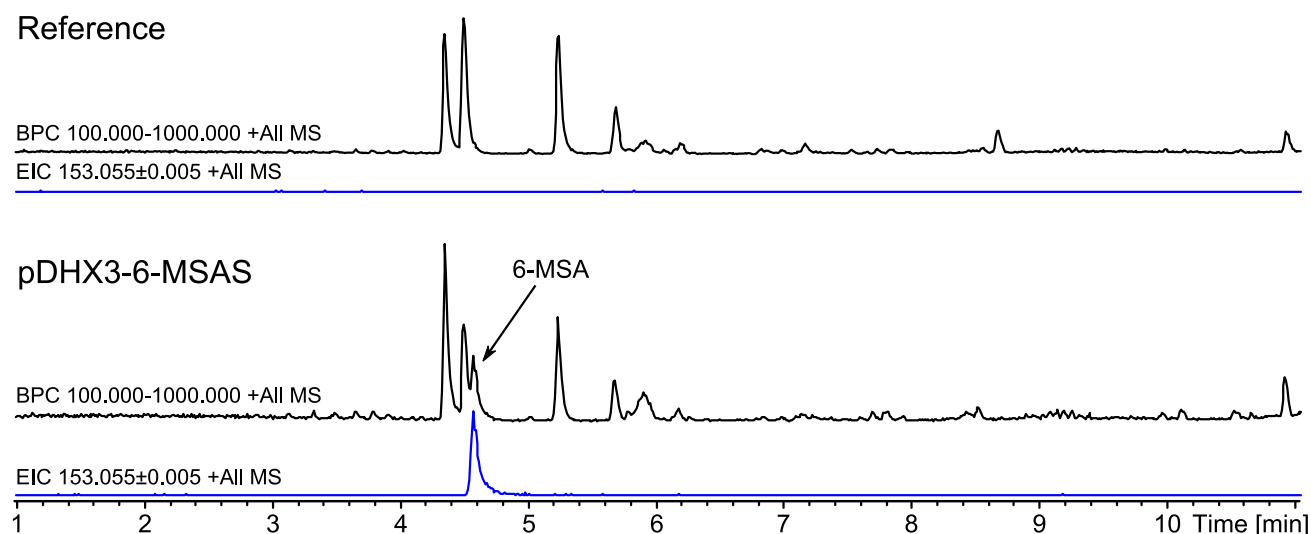


Figure 2. BPC and EIC (153.055 ± 0.005) for *A. aculeatus* reference and pDHX3-6-MSAS over-expression strain.

Over-expression of transcription factor from putative 6-MSA cluster results in up-regulation of three compounds. Examination of the gene cluster revealed that a putative transcription factor (TF) was located downstream of the 6-MSA synthase gene, why this could possibly be a regulator of the gene cluster. The TF was expressed from the expression platform pDHX2, cultivated on minimal and YES media, and analyzed by UHPLC-DAD-TOFMS. Examination of the metabolite profiles revealed up-regulation of two compounds on minimal medium, (**2**) and (**3**), which were not present in the reference (Figure 3).

The two compounds, (**2**) and (**3**), were produced by both the reference strain and pDHX2-TF when cultivated on YES, indicating that the biosynthesis of these compounds is normally active when

cultivated on YES medium, but not on minimal medium. In general, cultivation on YES medium did not result in a detectable change in the metabolite profile (Figure 3). An obvious explanation for this result is that the plasmid is lost when proper selection is not maintained; YES provides all nutrients, including uracil and uridine, whereas minimal medium does not. Another explanation is that the TF is already highly expressed during cultivation on YES, why over-expression does not affect production of the compounds.

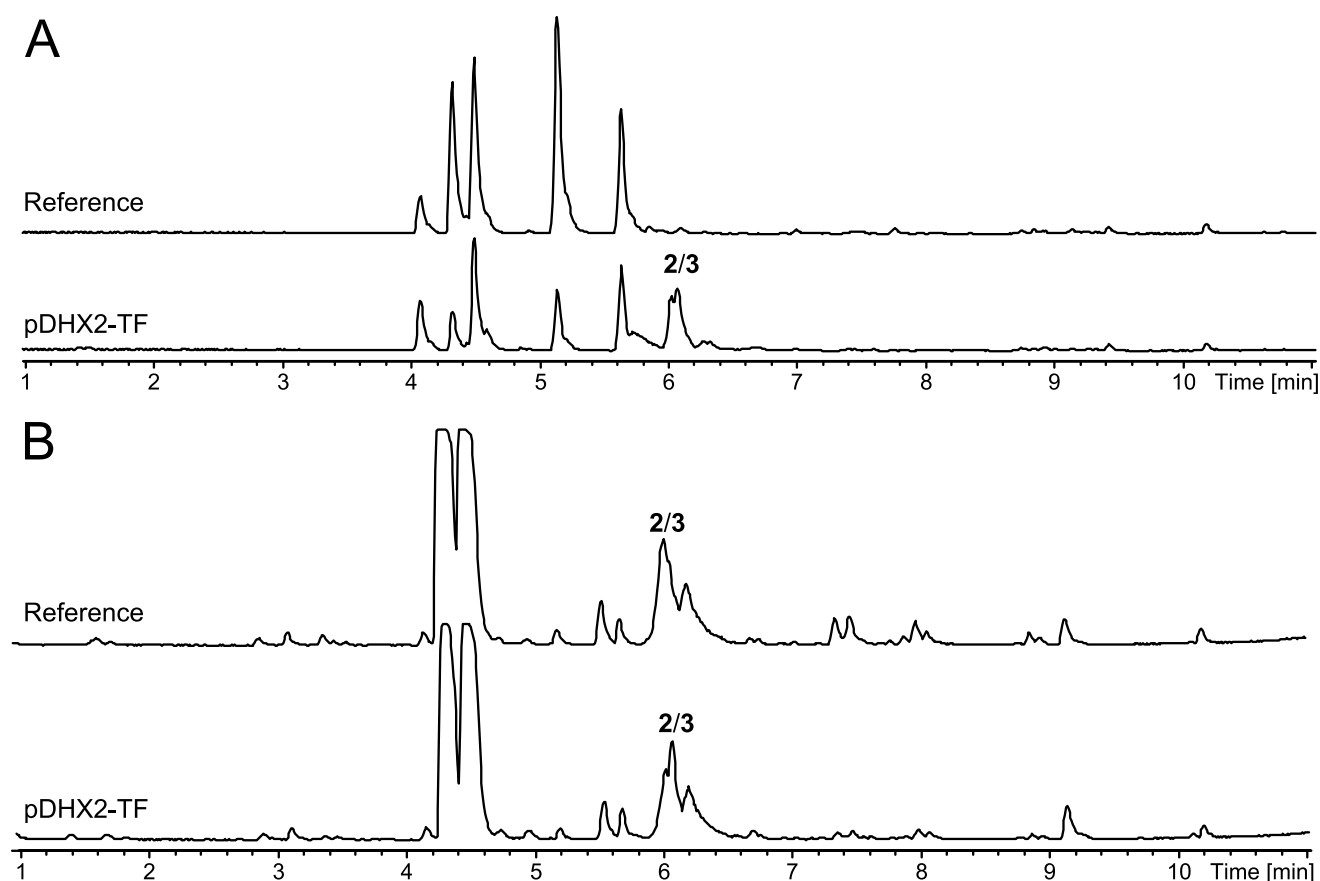


Figure 3. BPC of *A. aculeatus* expressing putative transcription factor from plasmid (pDHX2-TF) compared to a reference, cultivated on minimal medium (A) and YES (B).

Purification and structural elucidation reveals two novel, related compounds. Importantly, dereplication of the two compounds did not result in identification of any known compounds. (**2**) and (**3**) ($C_{21}H_{27}NO_5$) were isolated and structure elucidated by 2D NMR spectroscopy, revealing two analogous compounds, which have never been described before (Figure 4), NMR data can be found in supplementary Table S1 and Table S2.

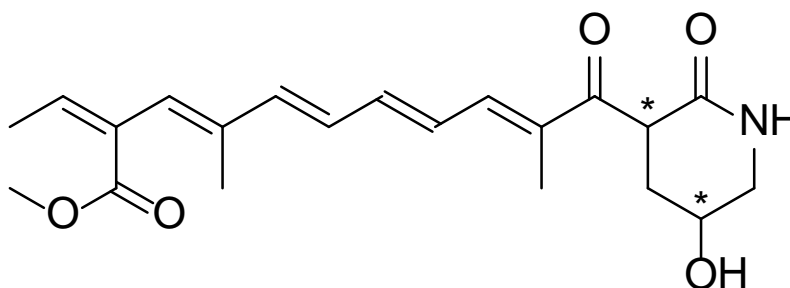


Figure 4. Structure of the novel compounds (**2**)/(**3**) only differing by their stereochemistry.

Stereocenters are marked with *.

The compounds (**2**) and (**3**) differ only by their stereochemistry. The structures of (**2**) and (**3**) are surprising, as it is not obvious how they are derived from 6-MSA, if at all. The backbone of the novel compounds have previously been observed e.g. in lucilactaene produced by a *Fusarium sp.* [9], epolactaene from a marine *Penicillium sp.* [10] and fusarin C from *Fusarium moniliforme* [11]. All of these compounds contain the long chain observed in compound (**2**) and (**3**) but differs in the attached moiety. Lucilactaene contains to fused five-membered rings, epolactaene a single five-membered lactam entailing an epoxide, and the fusarins display different structures. The novel compounds (**2**) and (**3**) are the first to entail a six-membered ring. The biosynthesis of the backbone chain has been investigated by labeling studies for the fusarins, which verified that it was of polyketide origin [12].

Over-expression of TF in *A. aculeatus* reveals novel compounds acurin A and B

Furthermore, labeling and genetic studies [13] showed that the final part originates from a C₄-unit. Assuming that (2) and (3) have the same polyketide chain, the final part must consist of a nitrogen-containing C₃ unit e.g. serine.

Fully labeled 6-MSA is incorporated into three compounds. To test whether 6-MSA was indeed incorporated into these compounds we fed fully labeled 6-MSA (¹³C₈-6-MSA) to the strain. Careful inspection of the metabolite profile revealed a mass shift of 7.023 Da for three compounds: (4), (5), and (6) (Figure 5).

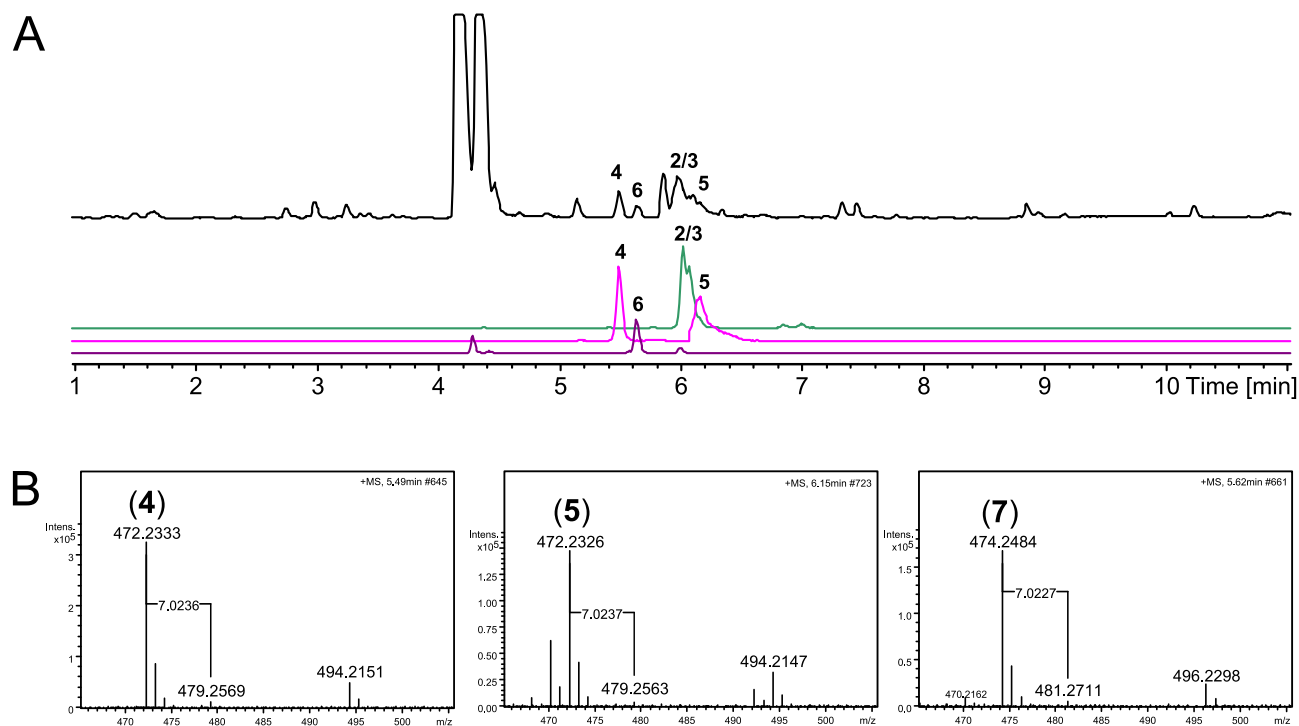


Figure 5. A. BPC (black) of pDHX2-TF on YES+¹³C₈-6-MSA, EIC 374.1963±0.005 (green), EIC 472.2333±0.005 (pink), EIC 474.2484±0.005 (purple). B. Mass spectra of compounds (4), (5), and (6).

Over-expression of TF in *A. aculeatus* reveals novel compounds acurin A and B

The fact that a mass shift equivalent of seven ^{13}C atoms, and not the eight ^{13}C atoms that are present in $^{13}\text{C}_8$ -6-MSA, suggests that one carbon atom is eliminated from 6-MSA during modification into the end products. A mass shift of 7.023 Da was also observed in feeding studies of the yanuthone D biosynthesis, where one carbon atom is eliminated from 6-MSA by decarboxylation [4]. Interestingly, (2) and (3), whose production appeared to be activated by expression of the TF, did not appear to incorporate 6-MSA (Figure S1). Based on the structures of (2) and (3), it is not surprising that 6-MSA is not incorporated into the compounds. In contrast, it is very surprising that a putative TF, located immediately downstream of a 6-MSA synthase, controls expression of two compounds, which does not depend on 6-MSA. Similarly, the three compounds that do contain 6-MSA, are not positively controlled by the putative TF in a simple manner, if at all. Thus, the mode of action for the putative TF is not clear.

At this point, we have not been able to produce compounds (4), (5), and (6) in large scale cultivations, why their structures remain unsolved.

Acurin A is not bioactive against cancer cells. Some of the compounds that are structurally similar to (2) and (3) have been reported as bioactive; Lucilactaene as a cell cycle inhibitor in p53-transfected cancer cell [9], epolactaene as a neuritogenic compound in human neuroblastoma cells [10], and fusarin C as a mutagen [11]. This prompted us to investigate the bioactivity of acurin A. The compound was tested in a CellTiter-Glo[®] assay [14] (Figure S2), but was found to be inactive.

ACKNOWLEDGMENTS

We kindly acknowledge Kenneth S. Bruno (Pacific Northwest National Laboratory, Richland, WA) for providing the KB1042 strain, and Kristian F. Nielsen (Technical University of Denmark) for providing fully labeled 6-MSA. This study was supported by grant 09-064967 from the Danish Council for Independent Research, Technology, and Production Sciences.

MATERIALS AND METHODS

Strains and media. The *A. nidulans* strain IBT29539 (*argB2*, *pyrG89*, *veA1*, *nkuAΔ*) [15] was used for heterologous expression of 6-MSAS. *A. aculeatus* ATCC16872-derived strain KB1042 (*pyr*⁻) was used as reference and for expression of 6-MSAS and TF. KB1042 was kindly provided by Kenneth S. Bruno (Pacific Northwest National Laboratory, Richland, WA, USA). *Escherichia coli* strain DH5α was used for cloning. Minimal media (MM) contained 1% glucose, 10 mM NaNO₃, 1× salt solution[16] and 2% agar for solid media. MM was supplemented with 10 mM uridine, 10 mM uracil, 4 mM L-arginine, 100 μg/ml hygromycin B (Invivogen, San Diego, CA, USA) when necessary.

Heterologous expression of 6-MSAS in *A. nidulans*. 6-MSAS from *A. aculeatus* ATCC16872 (ASPACDRAFT_1904397) was expressed heterologously in *A. nidulans*, using the expression platform developed in [8]. The 6-MSAS gene was PCR amplified using primers 6MSAS_ORF_FW (5'-AGAGCGAUATGCCTCACTTCTCCTCTGCC-3') and 6MSAS_ORF_RV (5'-

Over-expression of TF in *A. aculeatus* reveals novel compounds acurin A and B

TCTGCGAUGGTTAGAGATCTGATTAGCGGG-3'). Protoplasting, gene-targeting procedures and verification of strains were as described previously [4][17].

Expression of 6-MSAS and TF in *A. aculeatus*. 6-MSAS (ASPACDRAFT_1904397) and TF (ASPACDRAFT_123491) were expressed in *A. aculeatus*, using the AMA1-based expression platforms pDHX2 and pDHX3. pDHX2 is described in [4]. pDHX3 was constructed exactly as pDHX2, except *hph* was PCR amplified from pCB1003[18] using primers *hph*-pX-Fw (5'-ACTAGGTAAUGCTAGTGGAGGTCAACACATCA-3') and *hph*-pX-Rv (5'-AGTGGGGAUCGGTCGGCATCTACTCTATT -3') and inserted instead of *pyrG*. The 6-MSAS gene was PCR amplified as described above, and the TF was PCR amplified using primers TF_ORF_FW (5'-AGAGCGAUAAGCCATGTTCTCCACCTTCTC-3') and TF_ORF_RV (5'-TCTGCGAUTGTGACGGGACCCTAGATGAC-3'). Protoplasting and transformation procedures were carried out as described previously [17]. Transformants were selected on MM without uracil and uridine (for strains carrying pDHX2) or MM with hygromycin B (for strains carrying pDHX3).

Chemical analysis of strains. Unless otherwise stated strains were inoculated as a three point inoculation on solid MM media and incubated at 37 °C for 5 days. Extraction of metabolites was performed as described in [Smedsgaard, 1997]. 6-MSA was purchased from (Apin Chemicals, Oxon, UK). Analysis was performed using ultra-high-performance liquid chromatography (UPLC) UV/Vis diode array detector (DAD) high-resolution MS (TOFMS) on a maXis 3G orthogonal acceleration quadrupole time of flight mass spectrometer (Bruker Daltonics, Bremen, Germany) equipped with an

Over-expression of TF in *A. aculeatus* reveals novel compounds acurin A and B

electrospray ionization (ESI) source and connected to an Ultimate 3000 UHPLC system (Dionex, Sunnyvale, USA). The column used was a reverse-phase Kinetex 2.6- μmC_{18} , 100 mm \times 2.1mm (Phenomenex, Torrance, CA), and the column temperature was maintained at 40 °C. A linear water-acetonitrile (LCMS-grade) gradient was used (both solvents were buffered with 20 mM formic acid) starting from 10% (vol/vol) acetonitrile and increased to 100% in 10min, maintaining this rate for 3 min before returning to the starting conditions in 0.1 min and staying there for 2.4 min before the following run. A flow rate of 0.4 mL \cdot min⁻¹ was used. TOFMS was performed in ESI+ with a data acquisition range of 10 scans per second at m/z 100–1000. The TOFMS was calibrated using Bruker Daltonics high precision calibration algorithm (HPC) by means of the use of the internal standard sodium formate, which was automatically infused before each run. This provided a mass accuracy of better than 1.5 ppm in MS mode. UV/VIS spectra were collected at wavelengths from 200 to 700 nm. Data processing was performed using DataAnalysis 1.2 and Target Analysis4.0 software (Bruker Daltonics). Tandem MS was performed with fragmentation energy was varied from 18 to 55 eV.

Feeding experiments. Feeding experiments were carried out exactly as described in [4] on solid YES plates. Fully labeled 6-MSA (¹³C₈-6-MSA) was kindly provided by Kristian F. Nielsen (Technical University of Denmark), described in [4].

Preparative isolation of selected metabolites. *A. aculeatus* OeX-TF2 was cultivated on solid MM and incubated at 30°C for five days. The plates were harvested and extracted twice overnight with ethyl acetate (EtOAc) with 1 % formic acid. The extract was filtered and concentrated *in vacuo*. The

Over-expression of TF in *A. aculeatus* reveals novel compounds acurin A and B

combined extract was dissolved in methanol (MeOH) and H₂O (9:1), and an equal amount of heptane was added where after the phases were separated. To the MeOH/H₂O phase H₂O was added to a ratio of 1:1, and metabolites were then extracted with dichlormethane (DCM). The phases were then concentrated separately in vacuo.

The DCM phase was subjected to further purification on a semi-preparative HPLC, a Waters 600 Controller with a 996 photodiode array detector (Waters, Milford, MA, USA). This was achieved using a Luna II C₁₈ column (250 x 10 mm, 5 μ m, Phenomenex), a flowrate of 4 mL/min and a isocratic run at 40 % acetonitrile for 18 minutes. 50 ppm TFA was added to acetonitrile of HPLC grade and milliQ-water. This yielded 4.4 mg of acurin A and 5.0 mg of acurin B.

NMR and structural elucidation. The 1D and 2D spectra were recorded on a Unity Inova-500 MHz spectrometer (Varian, Palo Alto, CA, USA). Spectra were acquired using standard pulse sequences and ¹H spectra as well as DQF-COSY, NOESY, HSQC and HMBC spectra were acquired. The deuterated solvent was DMSO-*d*₆ and signals were referenced by solvent signals for DMSO-*d*₆ at $\delta_{\text{H}} = 2.49$ ppm and $\delta_{\text{C}} = 39.5$ ppm. The NMR data was processed in Bruker Topspin 3.1. Chemical shifts are reported in ppm (δ) and scalar couplings in hertz (Hz). The sizes of the *J* coupling constants reported in the tables are the experimentally measured values from the spectra. There are minor variations in the measurements which may be explained by the uncertainty of *J*. The structural elucidation can be found in supplementary and NMR data for acurin A and B compounds are listed in Tables S1 and S2, respectively.

REFERENCES

1. Hayashi, H.; Furutsuka, K.; Shiono, Y. Okaramines H and I, new okaramine congeners, from *Aspergillus aculeatus*. *Journal of natural products* **1999**, 62, 315–317.
2. Andersen, R.; Buechi, G.; Kobbe, B.; Demain, A. L. Secalonic acids D and F are toxic metabolites of *Aspergillus aculeatus*. *The Journal of Organic Chemistry* **1977**, 42, 352–353.
3. Petersen, L. M.; Hoeck, C.; Frisvad, J. C.; Gotfredsen, H., C.; Larsen, T. O. Dereplication guided biodiscovery of novel bioactives from the industrially important black fungus *Aspergillus aculeatus*. *Manuscript in preparation. Intended for Journal of Natural Products* **2013**.
4. Holm, D. K.; Petersen, L. M.; Klitgaard, A.; Knudsen, P. B.; Jarczynska, Z. D.; Nielsen, K. F.; Gotfredsen, C. H.; Larsen, T. O.; Mortensen, U. H. Molecular and chemical characterization of the biosynthesis of the 6-MSA derived meroterpenoid yanuthone D in *Aspergillus niger*. *JACS* **2013**, submitted.
5. Artigot, M. P.; Loiseau, N.; Laffitte, J.; Mas-Reguieg, L.; Tadrist, S.; Oswald, I. P.; Puel, O. Molecular cloning and functional characterization of two CYP619 cytochrome P450s involved in biosynthesis of patulin in *Aspergillus clavatus*. *Microbiology (Reading, England)* **2009**, 155, 1738–47.
6. Gems, D.; Johnstone, I. L.; Clutterbuck, A. J. An autonomously replicating plasmid transforms *Aspergillus nidulans* at high frequency. *Gene* **1991**, 98, 61–67.
7. Aleksenko, A.; Clutterbuck, A. J. The plasmid replicator AMA1 in *Aspergillus nidulans* is an inverted duplication of a low-copy-number dispersed genomic repeat. *Molecular microbiology* **1996**, 19, 565–574.
8. Hansen, B. G.; Salomonsen, B.; Nielsen, M. T.; Nielsen, J. B.; Hansen, N. B.; Nielsen, K. F.; Regueira, T. B.; Nielsen, J.; Patil, K. R.; Mortensen, U. H. Versatile enzyme expression and characterization system for *Aspergillus nidulans*, with the *Penicillium brevicompactum* polyketide synthase gene from the mycophenolic acid gene cluster as a test case. *Applied and environmental microbiology* **2011**, 77, 3044–3051.
9. Kakeya, H.; Kageyama, S.; Nie, L.; Onose, R.; Okada, G.; Beppu, T.; Norbury, C. J.; Osada, H. Lucilactaene, a new cell cycle inhibitor in p53-transfected cancer cells, produced by a *Fusarium* sp. *The Journal of antibiotics* **2001**, 54, 850–854.
10. Kakeya, H.; Takahashi, I.; Okada, G.; Isono, K.; Osada, H. Epolactaene, a novel neuritogenic compound in human neuroblastoma cells, produced by a marine fungus. *The Journal of antibiotics* **1995**, 48, 733–735.

11. Gelderblom, W.; Marasas, W.; Steyn, P.; PG, T.; Van der Merwe, K. J.; Rooyen, P.; Vleggaar, R.; Wessels, P. L. Structure elucidation of fusarin C, a mutagen produced by *Fusarium moniliforme*. *J. Chem. Soc., Chem. Commun.* **1984**, 122–124.
12. Steyn, P.; Vleggaar, R. Biosynthetic studies on the fusarins, metabolites of *Fusarium moniliforme*. *Journal of the Chemical Society, Chemical Communications* **1985**, 1189–1191.
13. Song, Z.; Cox, R. J.; Lazarus, C. M.; Simpson, T. J. Fusarin C biosynthesis in *Fusarium moniliforme* and *Fusarium venenatum*. *ChemBioChem* **2004**, 5, 1196–1203.
14. Bladt, T. T.; Dürr, C. S.; Knudsen, P. B.; Kildgaard, S.; Frisvad, J. C.; Gotfredsen, C. H.; Seiffert, M.; Larsen, T. O. Bio-activity guided discovery of ophiobolins with activity towards chronic lymphocytic leukemia. *Manuscript in preparation. Intended for molecules* **2013**.
15. Nielsen, J. B.; Nielsen, M. L.; Mortensen, U. H. Transient disruption of non-homologous end-joining facilitates targeted genome manipulations in the filamentous fungus *Aspergillus nidulans*. *Fungal Genetics and Biology* **2008**, 45, 165–170.
16. Cove, D. J. The induction and repression of nitrate reductase in the fungus *Aspergillus nidulans*. *Biochim Biophys Acta* **1966**, 113, 51–56.
17. Nielsen, M. L.; Albertsen, L.; Lettier, G.; Nielsen, J. B.; Mortensen, U. H. Efficient PCR-based gene targeting with a recyclable marker for *Aspergillus nidulans*. *Fungal genetics and biology* **2006**, 43, 54–64.
18. McCluskey, K.; Wiest, a.; Plamann, M. The Fungal Genetics Stock Center: a repository for 50 years of fungal genetics research. *Journal of Biosciences* **2010**, 35, 119–126.

Bibliography

- A. Aleksenko, D. Gems, and J. Clutterbuck. Multiple copies of mate elements support autonomous plasmid replication in *Aspergillus nidulans*. *Molecular Microbiology*, 20(2):427–434, 1996. ISSN 0950382x, 13652958.
- M. R. Andersen, M. P. Salazar, P. J. Schaap, P. J. I. van de Vondervoort, D. Culley, J. Thykaer, J. C. Frisvad, K. F. Nielsen, R. Albang, K. Albermann, R. M. Berka, G. H. Braus, S. A. Braus-Stromeyer, L. M. Corrochano, Z. Dai, P. W. M. van Dijck, G. Hofmann, L. L. Lasure, J. K. Magnuson, H. Menke, M. Meijer, S. L. Meijer, J. B. Nielsen, R. A. Samson, H. Stam, A. Tsang, J. M. van den Brink, A. Atkins, A. Aerts, H. Shapiro, J. Pangilinan, A. Salamov, Y. Lou, E. Lindquist, S. Lucas, J. Grimwood, I. V. Grigoriev, C. P. Kubicek, D. Martinez, N. N. M. E. van Peij, J. A. Roubos, J. Nielsen, and S. E. Baker. Comparative genomics of citric-acid-producing *Aspergillus niger* atcc 1015 versus enzyme-producing cbs 513.88. *Genome Research*, 21(6):885–897, 2011. ISSN 10889051, 15495469. doi: 10.1101/gr.112169.110.
- M. R. Andersen, J. B. Nielsen, A. Klitgaard, L. M. Petersen, M. Zachariasen, T. J. Hansen, L. H. Blicher, C. H. Gotfredsen, T. O. Larsen, K. F. Nielsen, and U. H. Mortensen. Accurate prediction of secondary metabolite gene clusters in filamentous fungi. *Proceedings of the National Academy of Sciences*, 110(1):20E99–E107, 2013. ISSN 00278424, 10916490. doi: 10.1073/pnas.1205532110.
- M. B. Austin and J. P. Noel. The chalcone synthase superfamily of type iii polyketide synthases. 2009.
- S. E. Baker. *Aspergillus niger* genomics: Past, present and into the future. *Medical Mycology*, 44(Suppl. 1):S17–S21, 2006. ISSN 13693786, 14602709.
- D. J. Ballance, F. P. Buxton, and G. Turner. Transformation of *Aspergillus nidulans* by the orotidine-5'-phosphate decarboxylase gene of *Neurospora crassa*. *Biochemical and Biophysical Research Communications*, 112(1):284–289, 1983. ISSN 0006291x, 10902104.
- S. Bergmann, J. Schümann, K. Scherlach, C. Lange, A. A. Brakhage, and C. Hertweck. Genomics-driven discovery of pks-nrps hybrid metabolites from *Aspergillus nidulans*. *Nature chemical biology*, 3(4):213–217, 2007. ISSN 15524450, 15524469. doi: 10.1038/nchembio869.

- F. R. Blattner, G. Plunkett, C. A. Bloch, N. T. Perna, V. Burland, M. Riley, J. ColladoVides, J. D. Glasner, C. K. Rode, G. F. Mayhew, J. Gregor, N. W. Davis, H. A. Kirkpatrick, M. A. Goeden, D. J. Rose, B. Mau, and Y. Shao. The complete genome sequence of *Escherichia coli* k-12. *Science*, 277(5331):1453–1474, 1997. ISSN 00368075, 10959203.
- M. Blumhoff, M. G. Steiger, H. Marx, D. Mattanovich, and M. Sauer. Six novel constitutive promoters for metabolic engineering of *Aspergillus niger*. *Applied Microbiology and Biotechnology*, 97(1):259–267, 2013. ISSN 01757598, 14320614. doi: 10.1007/s00253-012-4207-9.
- J. W. Bok, Y-M. Chiang, E. Szewczyk, Y. Reyes-Dominguez, A. D Davidson, J. F Sanchez, H-C. Lo, K. Watanabe, J. Strauss, B. R Oakley, C. C. C. Wang, and N. P. Keller. Chromatin-level regulation of biosynthetic gene clusters. *Nature chemical biology*, 5(7):462–464, 2009. ISSN 15524450, 15524469. doi: 10.1038/nchembio.177.
- D. W. Brown, J. H. Yu, H. S. Kelkar, M. Fernandes, T. C. Nesbitt, N. P. Keller, T. H. Adams, and T. J. Leonard. Twenty-five coregulated transcripts define a sterigmatocystin gene cluster in *Aspergillus nidulans*. *Proceedings of the National Academy of Sciences of the United States of America*, 93(4):1418–1422, 1996. ISSN 00278424, 10916490.
- F. P. Buxton, D. I. Gwynne, and R. W. Davies. Transformation of *Aspergillus niger* using the *argB* gene of *Aspergillus nidulans*. *Gene*, 37(1-3):207–214, 1985. ISSN 03781119, 18790038.
- Y-M. Chiang, E. Szewczyk, T. Nayak, A. D. Davidson, J. F. Sanchez, H-C. Lo, W-Y. Ho, H. Simityan, E. Kuo, A. Praseuth, K. Watanabe, B. R. Oakley, and C. C. C. Wang. Molecular genetic mining of the *Aspergillus* secondary metabolome: Discovery of the emericellamide biosynthetic pathway. *Chemistry & Biology*, 15(6):527–532, 2008. ISSN 10745521, 18791301. doi: 10.1016/j.chembiol.2008.05.010.
- Y-M. Chiang, E. Szewczyk, A. D. Davidson, N. Keller, B. R. Oakley, and C. C. C. Wang. A gene cluster containing two fungal polyketide synthases encodes the biosynthetic pathway for a polyketide, asperfuranone, in *Aspergillus nidulans*. *Journal of the American Chemical Society*, 131(8):2965–2970, 2009. ISSN 00027863, 15205126. doi: 10.1021/ja8088185.
- Y-M. Chiang, K. M. Meyer, M. Praseuth, S. E. Baker, K. S. Bruno, and C. C. C. Wang. Characterization of a polyketide synthase in *Aspergillus niger* whose product is a precursor for both dihydroxynaphthalene (dhn) melanin and naphtho- γ -pyrone. *Fungal Genetics and Biology*, 48(4):430–437, 2011. ISSN 10871845, 10960937. doi: 10.1016/j.fgb.2010.12.001.
- Y-M. Chiang, C. E. Oakley, M. Ahuja, R. Entwistle, A. Schultz, S-L. Chang, C. T. Sung, C. C. C. Wang, and B. R. Oakley. An efficient system for heterologous expression of secondary metabolite genes in *Aspergillus nidulans*. *Journal of the American Chemical Society*, 2013. ISSN 00027863, 15205126, 15205126. doi: 10.1021/ja401945a.

- A. J. Clutterbuck. Absence of laccase from yellow spored mutants of *Aspergillus nidulans*. *Journal of General Microbiology*, 70(3):423–435, 1972. ISSN 00221287.
- H. Conlon, I. Zadra, H. Haas, H. N. Arst, M. G. Jones, and M. X. Caddick. The *Aspergillus nidulans* gata transcription factor gene *areB* encodes at least three proteins and features three classes of mutation. *Molecular Microbiology*, 40(PART 2):361–375, 2001. ISSN 0950-382X, 0950382x, 13652958.
- R. J. Cox. Polyketides, proteins and genes in fungi: programmed nano-machines begin to reveal their secrets. *Organic & Biomolecular Chemistry*, 5(13):2010–2026, 2007. ISSN 14770520.
- R. J. Cox and T. J. Simpson. Fungal type i polyketide synthases. *Methods in Enzymology*, 459:49–78, 2009. ISSN 0076-6879, 00766879, 15577988.
- J. M. Crawford, B. C. R. Dancy, E. A. Hill, D. W. Udvary, and C. A. Townsend. Identification of a starter unit acyl-carrier protein transacylase domain in an iterative type i polyketide synthase. *Proceedings of the National Academy of Sciences of the United States of America*, 103(45):16728–16733, 2006. ISSN 00278424, 10916490. doi: 10.1073/pnas.0604112103.
- J. M. Crawford, P. M. Thomas, J. R. Scheerer, A. L. Vagstad, N. L. Kelleher, and C. A. Townsend. Deconstruction of iterative multidomain polyketide synthase function. *Science (New York, N.Y.)*, 320(5873):243–246, 2008. ISSN 00368075, 10959203. doi: 10.1126/science.1154711.
- J. M. Crawford, T. P. Korman, J. W. Labonte, A. L. Vagstad, E. A. Hill, O. Kamari-Bidkorpeh, S-C. Tsai, and C. A. Townsend. Structural basis for biosynthetic programming of fungal aromatic polyketide cyclization. *Nature*, 461(7267):1139–1143, 2009. ISSN 00280836.
- D. Croll and B. A. McDonald. Intron gains and losses in the evolution of *Fusarium* and *Cryptococcus* fungi. *Genome Biology and Evolution*, 4(11):201148–1161, 2012. ISSN 17596653. doi: 10.1093/gbe/evs091.
- D. Cullen, S. A. Leong, L. J. Wilson, and D. J. Henner. Transformation of *Aspergillus nidulans* with the hygromycin-resistance gene, *hph*. *Gene*, 57(1):21–26, 1987. ISSN 03781119, 18790038.
- E. H. Davidson, H. T. Jacobs, and R. J. Britten. Very short repeats and coordinate induction of genes. *Nature*, 301(5900):468–470, 1983. ISSN 00280836, 14764687.
- C. P. De Souza, S. B. Hashmi, and A. H. Osmani. Functional analysis of the *Aspergillus nidulans* kinome. *PLoS ONE*, 8(3):1, 2013. ISSN 19326203.
- S. Donadio, M. J. Staver, J. B. McAlpine, S. J. Swanson, and L. Katz. Modular organization of genes required for complex polyketide biosynthesis. *Science (New York, N.Y.)*, 252(5006):675–679, 1991. ISSN 00368075, 10959203.

- K. M. Fisch, A. F. Gillaspay, M. Gipson, J. C. Henrikson, A. R. Hoover, L. Jackson, F. Z. Najar, H. Wägele, and R. H. Cichewicz. Chemical induction of silent biosynthetic pathway transcription in *Aspergillus niger*. *Journal of Industrial Microbiology & Biotechnology*, 36(9):1199–1213, 2009. ISSN 13675435, 14765535. doi: 10.1007/s10295-009-0601-4.
- T. S. Fisher and V. A. Zakian. Ku: A multifunctional protein involved in telomere maintenance. *DNA Repair*, 4(11):1215–1226, 2005. ISSN 15687864.
- A. Fleming. *In-vitro* tests of penicillin potency. *The Lancet*, 239(6199):732–733, 1942. ISSN 01406736, 1474547x.
- A. Fleming. Streptococcal meningitis treated with penicillin - measurement of bacteriostatic power of blood and cerebrospinal fluid. *The Lancet*, 242(6267):434–438, 1943. ISSN 01406736, 1474547x.
- J. C. Frisvad, J. Smedsgaard, R. A. Samson, T. O. Larsen, and U. Thrane. Fumonisin b2 production by *Aspergillus niger*. *Journal of Agricultural and Food Chemistry*, 55(23):9727–9732, 2007. ISSN 00218561, 15205118. doi: 10.1021/jf0718906.
- J. C. Frisvad, T. O. Larsen, U. Thrane, M. Meijer, J. Varga, R. A. Samson, and K. F. Nielsen. Fumonisin and ochratoxin production in industrial *Aspergillus niger* strains. 2013.
- I. Fujii, Y. Ono, H. Tada, K. Gomi, Y. Ebizuka, and U. Sankawa. Cloning of the polyketide synthase gene *atX* from *Aspergillus terreus* and its identification as the 6-methylsalicylic acid synthase gene by heterologous expression. *Molecular and General Genetics*, 253(1-2):1–10, 1996. ISSN 00268925, 14321874. doi: 10.1007/s004380050289.
- J. E. Galagan, S. E. Calvo, C. Cuomo, L-J. Ma, J. R. Wortman, S. Batzoglou, S-I. Lee, M. Bastürkmen, C. C. Spevak, J. Clutterbuck, V. Kapitonov, J. Jurka, C. Scacciocchio, M. Farman, J. Butler, S. Purcell, S. Harris, G. H. Braus, O. Draht, S. Busch, C. D’Enfert, C. Bouchier, G. H. Goldman, D. Bell-Pedersen, S. Griffiths-Jones, J. H. Doonan, J. Yu, K. Vienken, A. Pain, M. Freitag, E. U. Selker, D. B. Archer, M. A. Peñalva, B. R. Oakley, M. Momany, T. Tanaka, T. Kumagai, K. Asai, M. Machida, W. C. Nierman, D. W. Denning, M. Caddick, M. Hynes, M. Paoletti, R. Fischer, B. Miller, P. Dyer, M. S. Sachs, S. A. Osmani, and B. W. Birren. Sequencing of *Aspergillus nidulans* and comparative analysis with *A. fumigatus* and *A. oryzae*. *Nature*, 438(7071):1105–1115, 2005. ISSN 00280836, 14764687. doi: 10.1038/nature04341.
- M. Ganzlin and U. Rinas. In-depth analysis of the *Aspergillus niger* glucoamylase (*glaA*) promoter performance using high-throughput screening and controlled bioreactor cultivation techniques. *Journal of Biotechnology*, 135(3):266–271, 2008. ISSN 01681656, 18734863. doi: 10.1016/j.jbiotec.2008.04.005.
- D. Gems, I. L. Johnstone, and A. J. Clutterbuck. An autonomously replicating plasmid transforms *Aspergillus nidulans* at high frequency. *Gene*, 98(1):61–67, 1991. ISSN 03781119, 18790038.

- R. Gerber, L. Lou, and L. Du. A plp-dependent polyketide chain releasing mechanism in the biosynthesis of mycotoxin fumonisins in *Fusarium verticillioides*. *Journal of the American Chemical Society*, 131(9):3148–3149, 2009. ISSN 00027863, 15205126. doi: 10.1021/ja8091054.
- F. Geu-Flores, H. H. Nour-Eldin, M. T. Nielsen, and B. A. Halkier. User fusion: a rapid and efficient method for simultaneous fusion and cloning of multiple pcr products. *Nucleic Acids Research*, 35(7):e55, 2007. ISSN 03051048.
- H. J. Gierman, M. H. G. Indemans, J. Koster, S. Goetze, J. Seppen, D. Geerts, R. van Driel, and R. Versteeg. Domain-wide regulation of gene expression in the human genome. *Genome Research*, 17(9):1286–1295, 2007. ISSN 10889051, 15495469. doi: 10.1101/gr.6276007.
- C. Gil Girol, K. M. Fisch, T. Heinekamp, S. Günther, W. Hüttel, J. Piel, A. A. Brakhage, and M. Müller. Regio- and stereoselective oxidative phenol coupling in *Aspergillus niger*. *Angewandte Chemie International Edition*, 51(39):9788–9791, 2012. ISSN 14337851, 15213773. doi: 10.1002/anie.201203603.
- A. Goffeau, B. G. Barrell, H. Bussey, R. W. Davis, B. Dujon, H. Feldmann, F. Galibert, J. D. Hoheisel, C. Jacq, M. Johnston, E. J. Louis, H. W. Mewes, Y. Murakami, P. Philippsen, H. Tettelin, and S. G. Oliver. Life with 6000 genes. *Science*, 274(5287):546–567, 1996. ISSN 00368075, 10959203.
- M. Goldoni, G. Azzalin, G. Macino, and C. Cogoni. Efficient gene silencing by expression of double stranded rna in *Neurospora crassa*. *Fungal Genetics and Biology*, 41(11):1016–1024, 2004. ISSN 10871845, 10960937. doi: 10.1016/j.fgb.2004.08.002.
- J-Y. Gou, X-H. Yu, and C-J. Liu. A hydroxycinnamoyltransferase responsible for synthesizing suberin aromatics in *Arabidopsis*. *Proceedings of the National Academy of Sciences of the United States of America*, 2009. ISSN 00278424, 10916490. doi: 10.1073/pnas.0905555106.
- B. G. Hansen, B. Salomonsen, M. T. Nielsen, J. B. Nielsen, N. B. Hansen, K. F. Nielsen, T. B. Regueira, J. Nielsen, K. R. Patil, and U. H. Mortensen. Versatile enzyme expression and characterization system for *Aspergillus nidulans*, with the *Penicillium brevicompactum* polyketide synthase gene from the mycophenolic acid gene cluster as a test case. *Applied and Environmental Microbiology*, 77(9):3044–3051, 2011. ISSN 00992240, 10985336. doi: 10.1128/AEM.01768-10.
- C. Henry, I. Mouyna, and J-P. Latgé. Testing the efficacy of rna interference constructs in *Aspergillus fumigatus*. *Current Genetics*, 51(4):277–284, 2007. ISSN 01728083, 14320983. doi: 10.1007/s00294-007-0119-0.
- P. Hernández-Ortiz and Eduardo A. Espeso. Phospho-regulation and nucleocytoplasmic trafficking of crza in response to calcium and alkaline-ph stress in *Aspergillus nidulans*. *Molecular Microbiology*, 89(3):532–551, 2013. ISSN 0950382x, 13652958. doi: 10.1111/mmi.12294.

- K. Ishiuchi, T. Nakazawa, T. Ookuma, S. Sugimoto, M. Sato, Y. Tsunematsu, N. Ishikawa, H. Noguchi, K. Hotta, H. Moriya, and K. Watanabe. Establishing a new methodology for genome mining and biosynthesis of polyketides and peptides through yeast molecular genetics. *ChemBioChem*, 13(6):846–854, 2012. ISSN 14394227, 14397633. doi: 10.1002/cbic.201100798.
- T. Itoh, K. Tokunaga, E. K. Radhakrishnan, I. Fujii, I. Abe, Y. Ebizuka, and T. Kushiro. Identification of a key prenyltransferase involved in biosynthesis of the most abundant fungal meroterpenoids derived from 3,5-dimethylorsellinic acid. *ChemBioChem*, 13(8):1132–1135, 2012. ISSN 14394227, 14397633. doi: 10.1002/cbic.201200124.
- F. Jacob and J. Monod. Genetic regulatory mechanisms in the synthesis of proteins. *Journal of molecular biology.*, 3:318–356, 1961. ISSN 00222836, 10898638.
- D. Jiang, Z. Hatahet, R. J. Melamede, Y. W. Kow, and S. S. Wallace. Characterization of escherichia coli endonuclease viii. *Journal of Biological Chemistry*, 272(51):32230–32239, 1997. ISSN 00219258, 1083351x.
- I. L. Johnstone, S. G. Hughes, and A. J. Clutterbuck. Cloning an *Aspergillus nidulans* developmental gene by transformation. *The EMBO journal*, 4(5):1307–1311, 1985. ISSN 02614189, 14602075.
- T. R. Jørgensen, J. Park, M. Arentshorst, A. M. van Welzen, G. Lamers, P. A. vanKuyk, R. A. Damveld, C. A. M. van den Hondel, K. F. Nielsen, J. C. Frisvad, and A. F. J. Ram. The molecular and genetic basis of conidial pigmentation in *Aspergillus niger*. *Fungal Genetics and Biology*, 48(5):544–553, 2011. ISSN 10871845, 10960937. doi: 10.1016/j.fgb.2011.01.005.
- L. Katz and S. Donadio. Polyketide synthesis: Prospects for hybrid antibiotics. *Annual Review of Microbiology*, 47:875–912, 1993. ISSN 00664227.
- J. T. Kealey, L. Liu, D. V. Santi, M. C. Betlach, and P. J. Barr. Production of a polyketide natural product in nonpolyketide-producing prokaryotic and eukaryotic hosts. *Proceedings of the National Academy of Sciences of the United States of America*, 95(2):505–509, 1998. ISSN 00278424, 10916490.
- N. P. Keller, G. Turner, and J. W. Bennett. Fungal secondary metabolism - from biochemistry to genomics. *Nature reviews. Microbiology*, 3(12):937–947, 2005. ISSN 17401526.
- D. E. Kelly, N. Kraševac, J. Mullins, and D. R. Nelson. The cypome (cytochrome p450 complement) of *Aspergillus nidulans*. *Fungal Genetics and Biology*, 46(1):5, 2009. ISSN 10871845, 10960937.
- J. Kennedy, K. Auclair, S. G. Kendrew, C. Park, J. C. Vederas, and C. R. Hutchinson. Modulation of polyketide synthase activity by accessory proteins during lovastatin biosynthesis. *Science (New York, N.Y.)*, 284(5418):1368–1372, 1999. ISSN 00368075, 10959203.
- N. Khaldi and K. H. Wolfe. Evolutionary origins of the fumonisin secondary metabolite gene cluster in *Fusarium verticillioides* and *Aspergillus niger*. *International journal of evolutionary biology.*, 2011:423821, 2011. ISSN 20908032, 2090052x.

- M. L. Klejstrup. *Mapping of secondary metabolites to their synthase genes in Aspergillus species*. DTU Systems Biology, 2012.
- S. Kroken, N. L. Glass, J. W. Taylor, O. C. Yoder, and B. G. Turgeon. Phylogenomic analysis of type i polyketide synthase genes in pathogenic and saprobic ascomycetes. *Proceedings of the National Academy of Sciences of the United States of America*, 100(26):15670–15675 and 3149064, 2003. ISSN 00278424.
- M. A. Larkin, G. Blackshields, N. P. Brown, R. Chenna, P. A. McGettigan, H. McWilliam, F. Valentin, I. M. Wallace, A. Wilm, R. Lopez, J. D. Thompson, T. J. Gibson, and D. G. Higgins. Clustal w and clustal x version 2.0. *Bioinformatics*, 23(21):2947–2948, 2007. ISSN 13674803, 14602059. doi: 10.1093/bioinformatics/btm404.
- D. W. Lee, M. Freitag, E. U. Selker, and R. Aramayo. A cytosine methyltransferase homologue is essential for sexual development in *Aspergillus nidulans*. *PLoS One*, 3(6):Article No.: e2531, 2008. ISSN 19326203. doi: 10.1371/journal.pone.0002531.
- É. Leiter, A. González, É. Erdei, C. Casado, L. Kovács, C. Ádám, J. Oláh, M. Miskei, M. Molnar, I. Farkas, Z. Hamari, J. Ariño, I. Pócsi, and V. Dombrádi. Protein phosphatase z modulates oxidative stress response in fungi. *Fungal Genetics and Biology*, 49(9):708–716, 2012. ISSN 10871845, 10960937. doi: 10.1016/j.fgb.2012.06.010.
- J. Li, Y. Luo, J-K. Lee, and H. Zhao. Cloning and characterization of a type iii polyketide synthase from *Aspergillus niger*. *Bioorganic & Medicinal Chemistry Letters*, 21(20):6085–6089, 2011a. ISSN 0960894x, 14643405. doi: 10.1016/j.bmcl.2011.08.058.
- Y. Li, Y-H. Chooi, Y. Sheng, J. S. Valentine, and Y. Tang. Comparative characterization of fungal anthracenone and naphthacenedione biosynthetic pathways reveals an α -hydroxylation-dependent claisen-like cyclization catalyzed by a dimanganese thioesterase. *Journal of the American Chemical Society*, 133(39):15773–15785, 2011b. ISSN 00027863, 15205126. doi: 10.1021/ja206906d.
- T. Lindahl. Dna repair enzymes. *Annual Review of Biochemistry*, 51:61–87, 1982. ISSN 00664154, 15454509. doi: 10.1146/annurev.bi.51.070182.000425.
- T. Lindahl, S. Ljungquist, W. Siebert, B. Nyberg, and B. Sperens. Dna n-glycosidases: properties of uracil-dna glycosidase from *Escherichia coli*. *The Journal of biological chemistry*., 252(10):3286–3294, 1977. ISSN 00219258, 1083351x.
- H-C. Lo, R. Entwistle, C-J. Guo, M. Ahuja, E. Szewczyk, J-H. Hung, Y-M. Chiang, B. R. Oakley, and C. C. C. Wang. Two separate gene clusters encode the biosynthetic pathway for the meroterpenoids austinol and dehydroaustinol in *Aspergillus nidulans*. *Journal of the American Chemical Society*, 134(10):4709–4720, 2012. ISSN 00027863, 15205126. doi: 10.1021/ja209809t.

- R. A. Lockington, H. M. Scaly-Lewis, C. Scazzocchio, and R. W. Davies. Cloning and characterization of the ethanol utilization regulon in *Aspergillus nidulans*. *Gene*, 33(2):137–149, 1985. ISSN 03781119, 18790038.
- P. F. Long, C. J. Wilkinson, C. P. Bisang, J. Cortes, N. Dunster, M. Oliynyk, E. McCormick, H. McArthur, C. Mendez, and J. A. Salas. Engineering specificity of starter unit selection by the erythromycin-producing polyketide synthase. *Molecular Microbiology*, 43(PART 5):1215–1226, 2002. ISSN 0950-382X, 0950382x, 13652958.
- A. Marchler-Bauer, S. Lu, J. B. Anderson, F. Chitsaz, M. K. Derbyshire, C. DeWeese-Scott, J. H. Fong, L. Y. Geer, R. C. Geer, N. R. Gonzales, M. Gwadz, D. I. Hurwitz, J. D. Jackson, Z. Ke, C. J. Lanczycki, F. Lu, G. H. Marchler, M. Mullokandov, M. V. Omelchenko, C. L. Robertson, J. S. Song, N. Thanki, R. A. Yamashita, D. Zhang, N. Zhang, C. Zheng, and S. H. Bryant. Cdd: a conserved domain database for the functional annotation of proteins. *Nucleic Acids Research*, 39(Suppl. 1):D225–D229, 2011. ISSN 03051048, 13624962. doi: 10.1093/nar/gkq1189.
- M. Markstein, C. Pitsouli, C. Villalta, S. E. Celniker, and N. Perrimon. Exploiting position effects and the gypsy retrovirus insulator to engineer precisely expressed transgenes. *Nature genetics*, 40(4):476–483, 2008. ISSN 10614036, 15461718. doi: 10.1038/ng.101.
- Y. Matsuda, T. Awakawa, T. Wakimoto, and I. Abe. Spiro-ring formation is catalyzed by a multifunctional dioxygenase in austinol biosynthesis. *Journal of the American Chemical Society*, 2013. ISSN 00027863, 15205126, 15205126. doi: 10.1021/ja405518u.
- K. McCluskey, A. Wiest, and M. Plamann. The fungal genetics stock center: a repository for 50 years of fungal genetics research. *Journal of Biosciences*, 35(1):119–126, 2010. ISSN 02505991, 09737138. doi: 10.1007/s12038-010-0014-6.
- R. J. Melamede, Z. Hatahet, Y. W. Kow, H. Ide, and S. W. Wallace. Isolation and characterization of endonuclease viii from *Escherichia coli*. *Biochemistry*, 33(5):1255–1264, 1994. ISSN 00062960, 15204995.
- V. Meyer, M. Arentshorst, A. El-Ghezal, A-C. Drews, R. Kooistra, C. A. M. J. J. van den Hondel, and A. F. J. Ram. Highly efficient gene targeting in the *Aspergillus niger kusA* mutant. *Journal of Biotechnology*, 128(4):770–775, 2007. ISSN 01681656, 18734863. doi: 10.1016/j.jbiotec.2006.12.021.
- V. Meyer, F. Wanka, J. van Gent, M. Arentshorst, C. A. M. J. J. van den Hondel, and A. F. J. Ram. Fungal gene expression on demand: An inducible, tunable, and metabolism-independent expression system for *Aspergillus niger*. *Applied and Environmental Microbiology*, 77(9):2975–2983, 2011. ISSN 00992240, 10985336. doi: 10.1128/AEM.02740-10.
- F-P. Miao, X-D. Li, X-H. Liu, R. H. Cichewicz, and N-Y. Ji. Secondary metabolites from an algicolous *Aspergillus versicolor* strain. *Marine Drugs*, 10(1):131–139, 2012. ISSN 16603397, 16603397. doi: 10.3390/md10010131.

- I. Molina, Y. Li-Beisson, F. Beisson, J. B. Ohlrogge, and M. Pollard. Identification of an *Arabidopsis* feruloyl-coenzyme a transferase required for suberin synthesis. *Plant Physiology*, 151(3):1317–1328, 2009. ISSN 00320889, 15322548. doi: 10.1104/pp.109.144907.
- F. J. Moralejo, R. E. Cardoza, S. Gutierrez, M. Lombrana, F. Fierro, and J. F. Martin. Silencing of the aspergillopepsin b (*pepB*) gene of *Aspergillus awamori* by antisense rna expression or protease removal by gene disruption results in a large increase in thaumatin production. *Applied and Environmental Microbiology*, 68(7):3550, 2002. ISSN 00992240, 10985336.
- T. Moriguchi, Y. Kezuka, T. Nonaka, Y. Ebizuka, and I. Fujii. Hidden function of catalytic domain in 6-methylsalicylic acid synthase for product release. *Journal of Biological Chemistry*, 285(20):15637–15643, 2010. ISSN 00219258, 1083351x. doi: 10.1074/jbc.M110.107391.
- N. R. Morris and A. P. Enos. Mitotic gold in a mold: *Aspergillus* genetics and the biology of mitosis. *Trends in Genetics*, 8(1):32–33, 1992. ISSN 01689525.
- H. Nakayashiki, S. Hanada, N. B. Quoc, N. Kadotani, Y. Tosa, and S. Mayama. Rna silencing as a tool for exploring gene function in ascomycete fungi. *Fungal Genetics and Biology*, 42(4):275–283, 2005. ISSN 10871845, 10960937. doi: 10.1016/j.fgb.2005.01.002.
- N. N. Nanthakumar, J. S. Dayton, and A. R. Means. Role of ca^{++} /calmodulin binding proteins in *Aspergillus nidulans* cell cycle regulation. *Progress in cell cycle research.*, 2:217–228, 1996. ISSN 10872957.
- T. Nayak, E. Szewczyk, C. E. Oakley, A. Osmani, L. Ukil, S. L. Murray, M. J. Hynes, S. A. Osmani, and B. R. Oakley. A versatile and efficient gene-targeting system for *Aspergillus nidulans*. *Genetics*, 172(3):1557, 2006. ISSN 00166731.
- D. J. Newman and G. M. Cragg. Natural products as sources of new drugs over the 30 years from 1981 to 2010. *Journal of Natural Products*, 75(3):311–335, 2012. ISSN 01633864, 15206025. doi: 10.1021/np200906s.
- C. B. Nielsen, B. Friedman, B. Birren, C. B. Burge, and J. E. Galagan. Patterns of intron gain and loss in fungi. *PLoS Biology*, 2(12):2234–2242, 2004. ISSN 15449173, 15457885.
- J. B. Nielsen, M. L. Nielsen, and U. H. Mortensen. Transient disruption of non-homologous end-joining facilitates targeted genome manipulations in the filamentous fungus *Aspergillus nidulans*. *Fungal Genetics and Biology*, 45(3):165–170, 2008. ISSN 10871845, 10960937. doi: 10.1016/j.fgb.2007.07.003.
- K. F. Nielsen, J. M. Mogensen, M. Johansen, T. O. Larsen, and J. C. Frisvad. Review of secondary metabolites and mycotoxins from the *Aspergillus niger* group. *Analytical and Bioanalytical Chemistry*, 395(5):1225–1242, 2009. ISSN 16182642, 16182650. doi: 10.1007/s00216-009-3081-5.
- M. L. Nielsen, L. Albertsen, G. Lettier, J. B. Nielsen, and U. H. Mortensen. Efficient pcr-based gene targeting with a recyclable marker for *Aspergillus nidulans*. *Fungal Genetics and Biology*, 43(1):54–64, 2006. ISSN 10871845.

- M. L. Nielsen, J. B. Nielsen, C. Rank, M. L. Klejnstrup, D. K. Holm, K. H. Brogaard, B. G. Hansen, J. C. Frisvad, T. O. Larsen, and U. H. Mortensen. A genome-wide polyketide synthase deletion library uncovers novel genetic links to polyketides and meroterpenoids in *Aspergillus nidulans*. *FEMS Microbiology Letters*, 321(2):157–166, 2011. ISSN 03781097, 15746968. doi: 10.1111/j.1574-6968.2011.02327.x.
- M. T. Nielsen, J. B. Nielsen, D. C. Anyaogu, D. K. Holm, K. F. Nielsen, T. O. Larsen, and U. H. Mortensen. Heterologous reconstitution of the intact geodin gene cluster in *Aspergillus nidulans* through a simple and versatile pcr based approach. *PLoS One*, 8(8), 2013. ISSN 19326203. doi: 10.1371/journal.pone.0072871.
- H. H. Nour-Eldin, B. G. Hansen, M. H. H. Nørholm, J. K. Jensen, and B. A. Halkier. Advancing uracil-excision based cloning towards an ideal technique for cloning pcr fragments. *Nucleic Acids Research*, 34(18):e122–e122, 2006. ISSN 03051048.
- M. H. H. Nørholm. A mutant pfu dna polymerase designed for advanced uracil-excision dna engineering. *BMC Biotechnology*, 10:21, 2010. ISSN 14726750. doi: 10.1186/1472-6750-10-21.
- C. E. Oakley, C. F. Weil, P. L. Kretz, and B. R. Oakley. Cloning of the *riboB* locus of *Aspergillus nidulans*. *Gene*, 53(2-3):293–298, 1987. ISSN 03781119, 18790038.
- L. R. Olsen, N. B. Hansen, M. T. Bonde, H. J. Genée, D. K. Holm, S. Carlsen, B. G. Hansen, K. R. Patil, U. H. Mortensen, and R. Wernersson. Phuser (primer help for user): a novel tool for user fusion primer design. *Nucleic acids research.*, 39(Web Server issue):W61–W67, 2011. ISSN 03051048, 13624962.
- J. M. Palmer, S. Mallareddy, D. W. Perry, J. F. Sanchez, J. M. Theisen, E. Szewczyk, B. R. Oakley, C. C. C. Wang, N. P. Keller, and P. M. Mirabito. Telomere position effect is regulated by heterochromatin-associated proteins and nkua in *Aspergillus nidulans*. *Microbiology (Reading)*, 156(Part 12):3522–3531, 2010. ISSN 13500872, 14652080. doi: 10.1099/mic.0.039255-0.
- H. J. Pel, J. H. de Winde, D. B. Archer, P. S. Dyer, G. Hofmann, P. J. Schaap, G. Turner, R. P. de Vries, R. Albang, K. Albermann, M. R. Andersen, J. D. Bendtsen, J. A. E. Benen, M. van den Berg, S. Breestraat, M. X. Caddick, R. Contreras, M. Cornell, P. M. Coutinho, E. G. J. Danchin, A. J. M. Debets, P. Dekker, P. W. M. van Dijk, A. van Dijk, L. Dijkhuizen, A. J. M. Driessen, C. d’Enfert, S. Geysens, C. Goosen, G. S. P. Groot, P. W. J. de Groot, T. Guillemette, B. Henrissat, M. Herweijer, J. P. T. W. van den Hombergh, C. A. M. J. J. van den Hondel, R. T. J. M. van der Heijden, R. M. van der Kaaij, F. M. Klis, H. J. Kools, C. P. Kubicek, P. A. van Kuyk, J. Lauber, X. Lu, M. J. E. C. van der Maarel, R. Meulenbergh, H. Menke, M. A. Mortimer, Jens. Nielsen, S. G. Oliver, M. Olsthoorn, K. Pal, N. N. M. E. van Peij, A. F. J. Ram, U. Rinas, J. A. Roubos, C. M. J. Sagt, M. Schmoll, J. Sun, D. Ussery, J. Varga, W. Vervecken, P. J. J. van de Vondervoort, H. Wedler, H. A. B. Wösten, A-P. Zeng, A. J. J. van Ooyen, J. Visser, and H. Stam. Genome sequencing and analysis of the versatile cell factory *Aspergillus niger* cbs 513.88. *Nature biotechnology*, 25(2):221–231, 2007. ISSN 10870156, 15461696. doi: 10.1038/nbt1282.

- J. H. Petersen. *Svamperiget*. Gads Forlag, second edition, 2003.
- D. J. Pitera, C. J. Paddon, J. D. Newman, and J. D. Keasling. Balancing a heterologous mevalonate pathway for improved isoprenoid production in *Escherichia coli*. *Metabolic Engineering*, 9(2):193–207, 2007. ISSN 10967176, 10967184. doi: 10.1016/j.ymben.2006.11.002.
- G. Pontecorvo, J. A. Roper, L. M. Hemmons, K. D. Macdonald, and A. W. J. Bufton. The genetics of *Aspergillus nidulans*. *Advances in genetics.*, 5:141–238, 1953. ISSN 00652660.
- R. H. Proctor, D. W. Brown, R. D. Plattner, and A. E. Desjardins. Co-expression of 15 contiguous genes delineates a fumonisin biosynthetic gene cluster in *Gibberella moniliformis*. *Fungal Genetics and Biology*, 38(2):237–249, 2003. ISSN 10871845, 10960937. doi: 10.1016/S1087-1845(02)00525-X.
- P. J. Punt, R. P. Oliver, M. A. Dingemanse, P. H. Pouwels, and C. A. M. J. J. van den Hondel. Transformation of *Aspergillus* based on the hygromycin b resistance marker from *Escherichia coli*. *Gene*, 56(1):117–124, 1987. ISSN 03781119, 18790038.
- P. J. Punt, M. A. Dingemanse, A. Kuyvenhoven, R. D. M. Soede, P. H. Pouwels, and C. A. M. J. J. van den Hondel. Functional elements in the promoter region of the *Aspergillus nidulans* *gpdA* gene encoding glyceraldehyde-3-phosphate dehydrogenase. *Gene*, 93(1):101–109, 1990. ISSN 03781119, 18790038.
- P. J. Punt, N. D. Zegers, M. Busscher, P. H. Pouwels, and C. A. M. J. J. van den Hondel. Intracellular and extracellular production of proteins in *Aspergillus* under the control of expression signals of the highly expressed *Aspergillus nidulans* *gpdA* gene. *Journal of Biotechnology*, 17(1):19–34, 1991. ISSN 01681656, 18734863. doi: 10.1016/0168-1656(91)90024-P.
- M. Rabache, J. Neumann, and J. Lavollay. Phenylpolyenes d’*Aspergillus niger*: structure et proprietes de l’aspermol. *Phytochemistry*, 13(3):637–642, 1974. ISSN 00319422, 18733700.
- C. Rasmussen, C. Garen, S. Brining, R. L. Kincaid, R. L. Means, and A. R. Means. The calmodulin-dependent protein phosphatase catalytic subunit (calcineurin a) is an essential gene in *Aspergillus nidulans*. *The EMBO journal*, 13(16):3917–3924, 1994. ISSN 02614189, 14602075.
- A. H. F. J. Roth and P. Dersch. A novel expression system for intracellular production and purification of recombinant affinity-tagged proteins in *Aspergillus niger*. *Applied Microbiology and Biotechnology*, 86(2): 659–670, 2010. ISSN 01757598, 14320614. doi: 10.1007/s00253-009-2252-9.
- M. Röttig, M. H. Medema, K. Blin, T. Weber, C. Rausch, and O. Kohlbacher. Nrpspredictor2—a web server for predicting nrps adenylation domain specificity. *Nucleic acids research.*, 39(Web Server issue):W362–W367, 2011. ISSN 03051048, 13624962. doi: 10.1093/nar/gkr323.
- P. Rugbjerg, M. Naesby, and U. H. Mortensen. Reconstruction of the biosynthetic pathway for the core fungal polyketide scaffold rubrofusarin in *Saccharomyces cerevisiae*. *Microbial Cell Factories*, 12(1):1, 2013. ISSN 14752859.

- R. A. Samson, J. Houbroken, U. Thrane, J. C. Frisvad, and B. Andersen. Food and indoor fungi.
- R. A. Samson, J. Houbroken, U. Thrane, J. C. Frisvad, and B. Andersen. *Food and Indoor Fungi*. CBS-KNAW, first edition, 2010.
- V. Schroeckh, K. Scherlach, H-W. Nützmann, E. Shelest, W. Schmidt-Heck, J. Schuemann, K. Martin, C. Hertweck, and A. A. Brakhage. Intimate bacterial-fungal interaction triggers biosynthesis of archetypal polyketides in *Aspergillus nidulans*. *Proceedings of the National Academy of Sciences of the United States of America*, 106(34):14558–14563, 2009. ISSN 00278424.
- J. Schumann and C. Hertweck. Advances in cloning, functional analysis and heterologous expression of fungal polyketide synthase genes. *Journal of Biotechnology*, 124(4):690, 2006. ISSN 01681656.
- J. Smedsgaard. Micro-scale extraction procedure for standardized screening of fungal metabolite production in cultures. *Journal of Chromatography A*, 760(2):264–270, 1997. ISSN 00219673, 18733778.
- S. Son and S. A. Osmani. Analysis of all protein phosphatase genes in *Aspergillus nidulans* identifies a new mitotic regulator, fcp1. *Eukaryotic Cell*, 8(4):573–585, 2009. ISSN 15359778, 15359786. doi: 10.1128/EC.00346-08.
- R. Sopko, D. Huang, N. Preston, G. Chua, B. Papp, K. Kafadar, M. Snyder, S. G. Oliver, M. Cyert, T. R. Hughes, C. Boone, and B. Andrews. Mapping pathways and phenotypes by systematic gene overexpression. *Molecular Cell*, 21(3):319–330, 2006. ISSN 10972765, 10974164.
- C. Sousa, V. deLorenzo, and A. Cebolla. Modulation of gene expression through chromosomal positioning in *Escherichia coli*. *Microbiology-UK*, 143:2071–2078, 1997. ISSN 13500872, 14652080.
- M. Stanke and B. Morgenstern. Augustus: a web server for gene prediction in eukaryotes that allows user-defined constraints. *Nucleic Acids Research*, 33(Suppl. S):W465–W467, 2005. ISSN 03051048, 13624962.
- M. Stanke, R. Steinkamp, S. Waack, and B. Morgenstern. Augustus: a web server for gene finding in eukaryotes. *Nucleic acids research.*, 32(Web Server issue):W309–W312, 2004. ISSN 03051048, 13624962.
- E. Szewczyk, Y-M. Chiang, C. E. Oakley, A. D. Davidson, C. C. C. Wang, and B. R. Oakley. Identification and characterization of the asperthecin gene cluster of *Aspergillus nidulans*. *Applied and Environmental Microbiology*, 74(24):7607–7612, 2008. ISSN 00992240, 10985336. doi: 10.1128/AEM.01743-08.
- A. Thompson and M. J. Gasson. Location effects of a reporter gene on expression levels and on native protein synthesis in *Lactococcus lactis* and *Saccharomyces cerevisiae*. *Applied and Environmental Microbiology*, 67(PART 8):3434–3439, 2001. ISSN 0099-2240, 00992240, 10985336.
- J. Tilburn, C. Scazzocchio, G. G. Taylor, J. H. Zabicky-Zissman, R. A. Lockington, and R. W. Davies. Transformation by integration in *Aspergillus nidulans*. *Gene*, 26(2-3):205–221, 1983. ISSN 03781119, 18790038.

- S. F.F. Torriani, E. H. Stukenbrock, P. C. Brunner, B. A. McDonald, and D. Croll. Evidence for extensive recent intron transposition in closely related fungi. *Current Biology*, 21(23):2017–2022, 2011. ISSN 09609822, 18790445. doi: 10.1016/j.cub.2011.10.041.
- J. Varga, K. Rigo, B. Toth, J. Teren, and Z. Kozakiewicz. Evolutionary relationships among *Aspergillus* species producing economically important mycotoxins. *Food Technology and Biotechnology*, 41(1):29–36, 2003. ISSN 13309862.
- J. C. Verdoes, P. J. Punt, and C. A. M. J. J. van den Hondel. Molecular genetic strain improvement for the overproduction of fungal proteins by filamentous fungi. *Applied Microbiology and Biotechnology*, 43(2): 195–205, 1995. ISSN 01757598, 14320614. doi: 10.1007/BF00172812.
- A. Watanabe, Y. Ono, I. Fujii, U. Sankawa, M. E. Mayorga, W. E. Timberlake, and Y. Ebizuka. Product identification of polyketide synthase coded by *Aspergillus nidulans* *wA* gene. *Tetrahedron Letters*, 39(42): 7733–7736, 1998. ISSN 00404039, 18733581.
- A. Watanabe, I. Fujii, U. Sankawa, M. E. Mayorga, W. E. Timberlake, and Y. Ebizuka. Re-identification of *Aspergillus nidulans* *wA* gene to code for a polyketide synthase of naphthopyrone. *Tetrahedron Letters*, 40(1):91–94, 1999. ISSN 00404039, 18733581. doi: 10.1016/S0040-4039(98)80027-0.
- K. J. Weissman. Biochemistry. anatomy of a fungal polyketide synthase. *Science (New York, N.Y.)*, 320(5873):186–187, 2008. ISSN 00368075.
- K. H. Wong, M. J. Hynes, R. B. Todd, and M. A. Davis. Deletion and overexpression of the *Aspergillus nidulans* gata factor *areb* reveals unexpected pleiotropy. *Microbiology (Reading)*, 155(Part 12, Sp. Iss. SI): 3868–3880, 2009. ISSN 13500872, 14652080. doi: 10.1099/mic.0.031252-0.
- J. R. Wortman, J. M. Gilsenan, V. Joardar, J. Deegan, J. Clutterbuck, M. R. Andersen, D. Archer, M. Bencina, G. Braus, P. Coutinho, H. von Döhren, J. Doonan, A. J. M. Driessen, P. Durek, E. Espeso, E. Fekete, M. Flipphi, C. G. Estrada, S. Geysens, G. Goldman, P. W. J. de Groot, K. Hansen, S. D. Harris, T. Heinekamp, K. Helmstaedt, B. Henrissat, G. Hofmann, T. Homan, T. Horio, H. Horiuchi, S. James, M. Jones, L. Karaffa, Z. Karányi, M. Kato, N. Keller, D. E. Kelly, J. A. K. W. Kiel, J.-M. Kim, I. J. van der Klei, F. M. Klis, A. Kovalchuk, N. Kraševac, C. P. Kubicek, B. Liu, A. MacCabe, V. Meyer, P. Mirabito, M. Miskei, M. Mos, J. Mullins, D. R. Nelson, J. Nielsen, B. R. Oakley, S. A. Osmani, T. Pakula, A. Paszewski, I. Paulsen, S. Pilsyk, I. Pócsi, P. J. Punt, A. F. J. Ram, Q. Ren, X. Robellet, G. Robson, B. Seiboth, P. van Solingen, T. Specht, J. Sun, N. Taheri-Talesh, N. Takeshita, D. Ussery, P. A. vanKuyk, H. Visser, P. J. I. van de Vondervoort, R. P. de Vries, J. Walton, X. Xiang, Y. Xiong, A. P. Zeng, B. W. Brandt, M. J. Cornell, C. A. M. J. J. van den Hondel, J. Visser, S. G. Oliver, and G. Turner. The 2008 update of the *Aspergillus nidulans* genome annotation: A community effort. *Fungal Genetics and Biology*, 46(1):2, 2009. ISSN 10871845, 10960937.

- J. Yaegashi, M. B. Praseuth, S-W. Tyan, J. F. Sanchez, R. Entwistle, Y-M. Chiang, B. R. Oakley, and C. C. C. Wang. Molecular genetic characterization of the biosynthesis cluster of a prenylated isoindolinone alkaloid aspernidine a in *Aspergillus nidulans*. *Organic Letters*, 15(11):202862–2865, 2013. ISSN 15237060, 15237052, 15237052. doi: 10.1021/ol401187b.
- S. Yamane, M. Yamaoka, M. Yamamoto, T. Maruki, H. Matsuzaki, T. Hatano, and S. Fukui. Region specificity of chromosome iii on gene expression in the yeast *Saccharomyces cerevisiae*. *Journal of General and Applied Microbiology*, 44(4):275–281, 1998. ISSN 00221260, 13498037.
- M. M. Yelton, J. E. Hamer, and W. E. Timberlake. Transformation of *Aspergillus nidulans* by using a *trpC* plasmid. *Proceedings of the National Academy of Sciences of the United States of America*, 81(5):1470–1474, 1984. ISSN 00278424, 10916490.
- J-H. Yu and T. J. Leonard. Sterigmatocystin biosynthesis in *Aspergillus nidulans* requires a novel type i polyketide synthase. *Journal of Bacteriology*, 177(16):4792–4800, 1995. ISSN 00219193, 10985530.
- A. O. Zabala, W. Xu, Y-H. Chooi, and Y. Tang. Characterization of a silent azaphilone gene cluster from *Aspergillus niger* atcc 1015 reveals a hydroxylation-mediated pyran-ring formation. *Chemistry and Biology*, 19(8):1049–1059, 2012.
- L-Y. Zhang, Y-F. Yang, and D-K. Niu. Evaluation of models of the mechanisms underlying intron loss and gain in *Aspergillus* fungi. *Journal of Molecular Evolution*, 71(5-6):364–373, 2010. ISSN 00222844, 14321432. doi: 10.1007/s00239-010-9391-6.
- X. F. Zheng, Y. Kobayashi, and M. Takeuchi. Construction of a low-serine-type-carboxypeptidase-producing mutant of *Aspergillus oryzae* by the expression of antisense rna and its use as a host for heterologous protein secretion. *Applied Microbiology and Biotechnology*, 49(1):39–44, 1998. ISSN 01757598, 14320614. doi: 10.1007/s002530051134.
- H. Zhou, K. Qiao, Z. Gao, J. C. Vederas, and Y. Tang. Insights into radicicol biosynthesis via heterologous synthesis of intermediates and analogs. *Journal of Biological Chemistry*, 285(53):41412–41421, 2010. ISSN 00219258, 1083351x. doi: 10.1074/jbc.M110.183574.
- H. Zhou, Z. Gao, K. Qiao, J. Wang, J. C. Vederas, and Y. Tang. A fungal ketoreductase domain that displays substrate-dependent stereospecificity. *Nature chemical biology*, 8(4):331–333, 2012. ISSN 15524450, 15524469. doi: 10.1038/nchembio.912.
- A. Zuccaro, S. Götze, S. Kneip, P. Dersch, and J. Seibel. Tailor-made fructooligosaccharides by a combination of substrate and genetic engineering. *ChemBioChem*, 9(1):143–149, 2008. ISSN 14394227, 14397633. doi: 10.1002/cbic.200700486.

Appendices

Appendix A-H, pp. 270-398

Appendix A

Supporting information for Chapter 4

A Genome-wide Polyketide Synthase Deletion Library uncovers Novel Genetic Links to Polyketides and Meroterpenoids in *Aspergillus nidulans*

Michael L. Nielsen, Jakob B. Nielsen, Christian Rank, Marie L. Klejnstrup, Dorte M. K. Holm, Katrine H. Brogaard, Bjarne G. Hansen, Jens C. Frisvad, Thomas O. Larsen and Uffe H. Mortensen

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Figures (page 3)

Fig. S1. Eight known metabolites that have been linked to specific PKS genes in *A. nidulans*.

Fig. S2. A graphical illustration of the procedure employed to make the gene targeting fragments for the PKS deletions.

Fig. S3. Procedure for diagnostic PCR.

Fig. S4. Chromosome map showing the position of the 32 PKS genes.

Fig. S5. Verification that the polyketide is absent in selected deletion mutants.

Fig. S6. Positive mode extracted ion chromatograms for the *mdpGΔ* strain cultivated on RTO.

Fig. S7. Additional polyketides that were detected in metabolite extracts in this study.

Fig. S8. The mutants *orsAΔ* and AN7903Δ lack production of violaceols.

Fig. S9. Extracted ion chromatograms of metabolic extracts from the reference and *ausAΔ* strains.

Fig. S10. ¹H NMR spectrum of 3,5-dimethyl orsellinic acid in DMSO-*d*₆.

Fig. S11. ¹H NMR spectrum of dehydroaustinol in CDCl₃.

Fig. S12. ¹H NMR spectrum of arugosin A open form in DMSO-*d*₆.

Supporting Information

Text (page 15)

Text S1. Details about compound identification.

Tables (page 17)

Table S1. PCR primers for Gateway assembly.

Table S2. Oligonucleotide primers for diagnostic PCR.

Table S3. Additional oligonucleotides used in the study.

Table S4. The constructed *A. nidulans* strains have been deposited into the IBT Culture Collection.

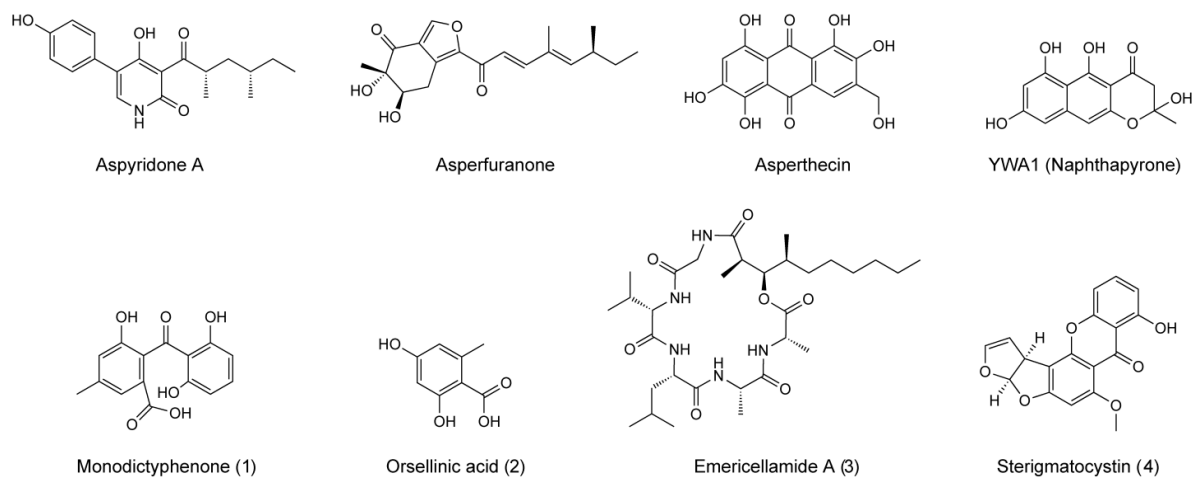
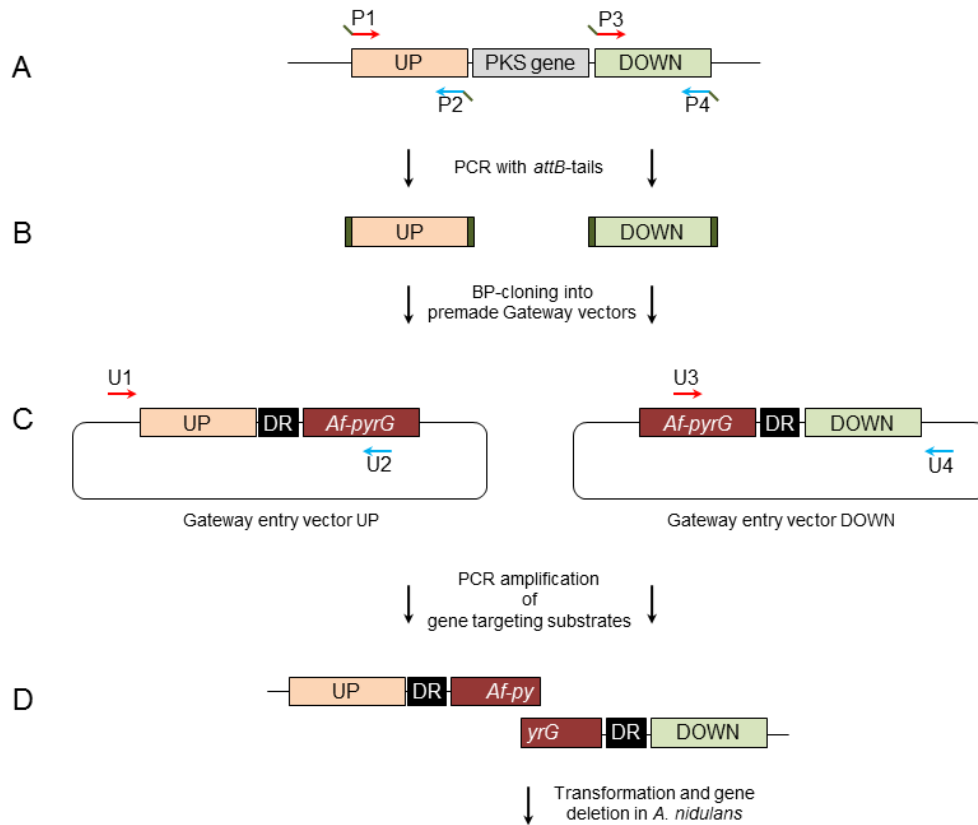


Fig. S1. Eight known metabolites that have been linked to specific PKS genes in *A. nidulans*.

Aspyridone (*adpA*), asperfuranone (*afoE/G*), asperthecin (*aptA*), YWA1 (*wA*), (1) monodictyphenone (*mdpG*), (2) orsellinic acid (*orsA*), (3) emericellamide (A) (*easB*) and (4) sterigmatocystin (*stcA*).

8



9

Fig. S2. A graphical illustration of the procedure employed to make the gene targeting fragments for the PKS deletions. Genomic DNA serves as template for two separate PCR reactions (A) that amplify the UP-stream and DOWN-stream regions of the PKS gene targeted for deletion. The oligonucleotide pairs contain approx. 20 nucleotide *attB* sequences that allow BP-recombination according to the Gateway cloning system (Invitrogen). (B) The PCR products with the specific *attB* sequences are incorporated into donor vectors (C) that contain *A. fumigatus pyrG* flanked by a sequence (DR) that serves as a direct repeat in the final mutant. The resulting plasmids (D) are purified from *E. coli*, and used as template to amplify gene-targeting substrates by PCR. Universal oligonucleotides U1 & U2 are used to amplify the UP-stream targeting fragments; and D1 and D2 are used to amplify DOWN-stream targeting fragments. The final targeting substrates (E) are co-transformed as bipartite gene-targeting substrates.

21

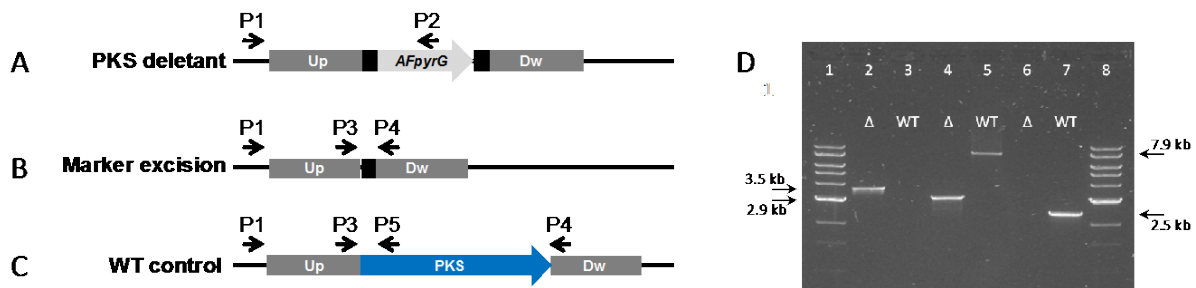


Fig. S3. Illustration of the diagnostic PCR procedure. A standard repertoire of PCR reactions were used in the analysis of PKS deletion mutant strains. (A) shows PCR check of a targeting event for a given PKS locus. Up and Dw depicts the up- and downstream flanks of the PKS, respectively. The forward primer, P1, anneals outside the targeting sequence and the reverse primer, P2, is placed within the *A. fumigatus pyrG* (*AFpyrG*) marker gene. (B) the marker gene has been excised from the locus and the primer pair, P1 or P3 with P4, binds on either side of the remaining direct repeat sequence. The reaction can also occur on reference strain genomic DNA template, but in this case to produce a significantly larger fragment as illustrated in (C). In addition to this, a PCR reaction (C) using P1 or P3 with P5 positioned inside the PKS gene, identifies false positives. The PKS gene is shown as a blue arrow, the marker gene as a light grey arrow and the direct repeat as black squares. Drawing is not to scale. A control PCR reaction with reference strain genomic DNA as template is always included. (D) shows a typical agarose gel result for the PCR verification of a deletion strain (AN0150 (*mdpG*) in this case). Lanes are indicated by numbers 1-8 below the wells. Lane 1 & 8, 1 kb ladder from New England Biolabs. Lane 2 shows the test for correct replacement of *mdpG* with the *AFpyrG* marker by P1 & P2 (test A) in the deletion strain candidate giving the 3.5 kb band, and the expected absence of a band in the WT control reaction in lane 3. Lanes 4 and 5, confirms marker loss after counter-selecting on 5-FOA via primers P1 and P4 (test B), the pop-out event is verified by the 2.9 kb band in lane 4 whereas WT gave the expected larger 7.9 kb band. Lane 6 and 7 with primers P1 & P5 confirm that the deletion strain is pure and does not contain WT nuclei (intact *mdpG* sequence). All spore PCR reactions were run in triplicates.

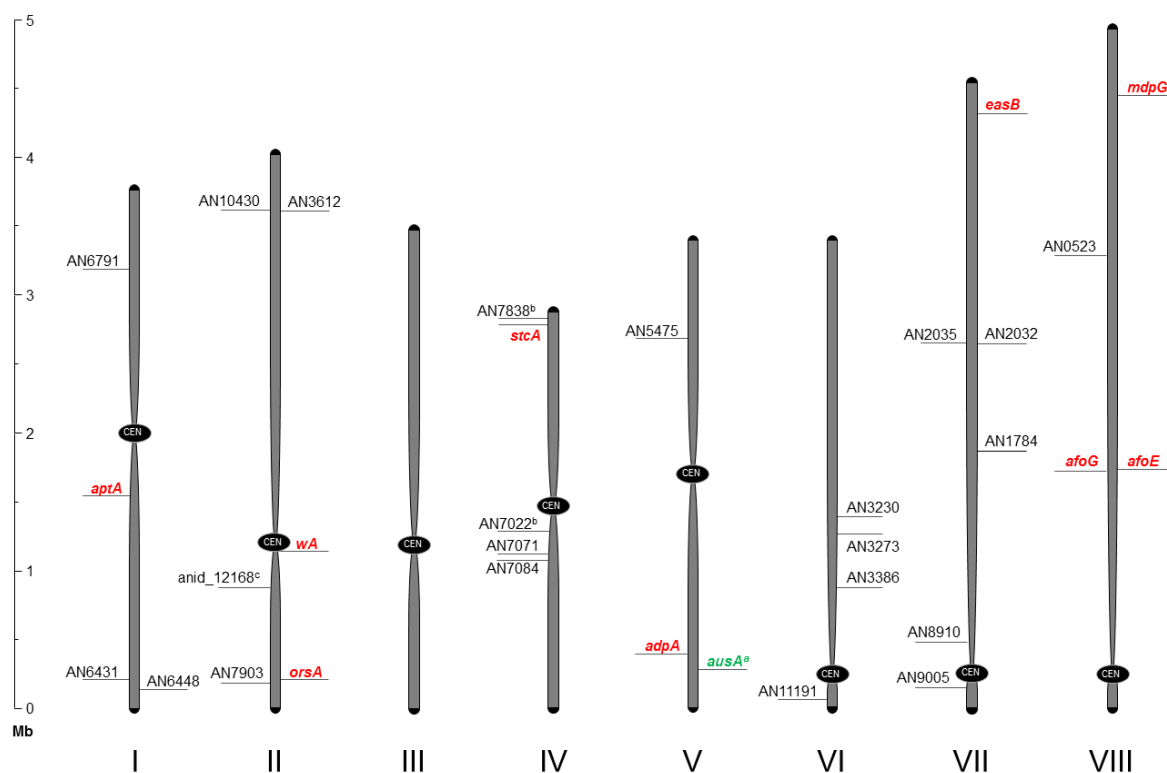
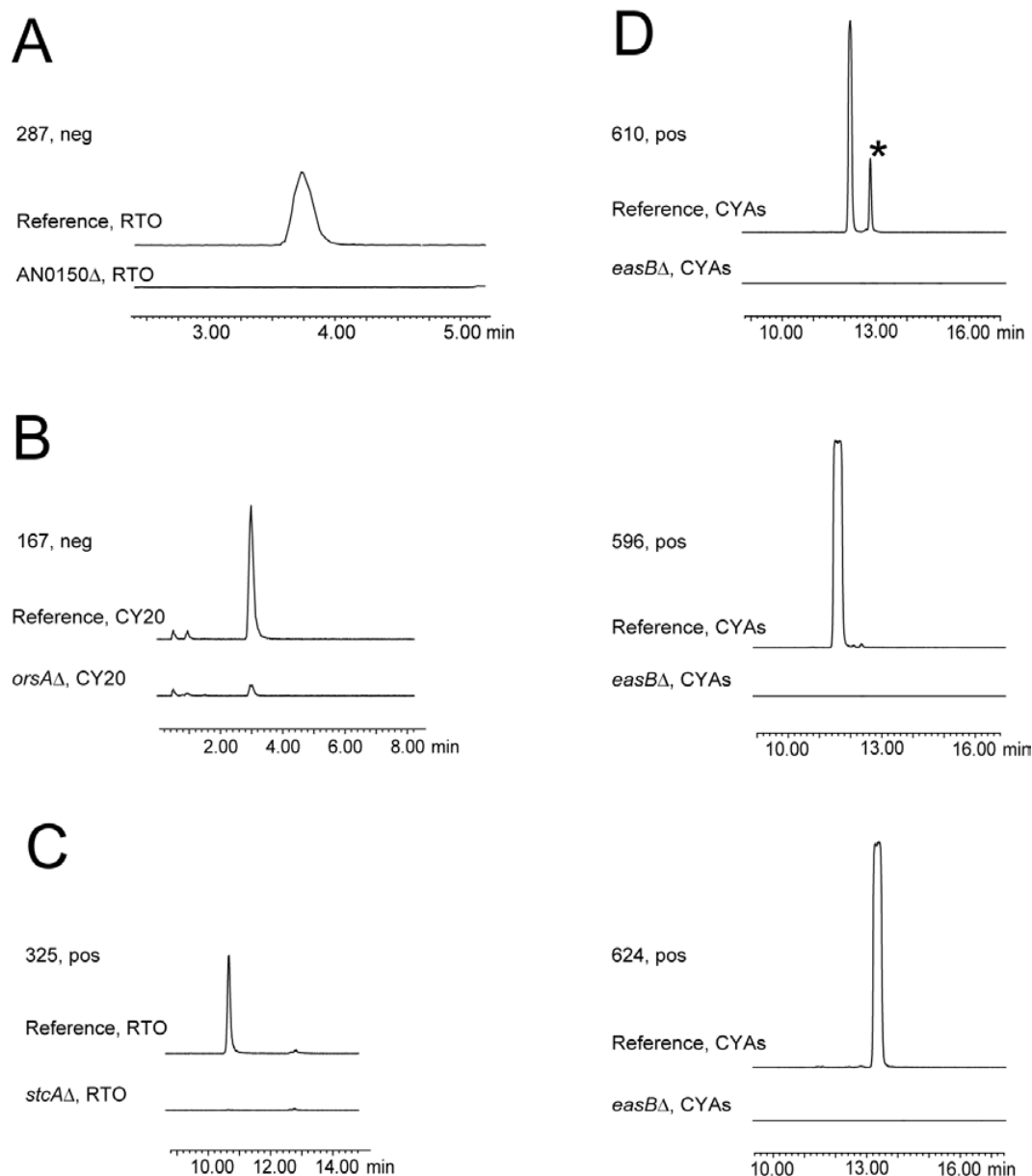


Fig. S4. Chromosome map showing the position of the 32 PKS genes on the eight chromosomes of *A. nidulans*. The map was generated based mainly on information from the *Aspergillus* Genome Database at Stanford. The approximate location of the centromere is shown for each of the eight chromosomes. Telomeres are indicated as black tips and are not to scale. PKS genes that have previously been linked to polyketide compounds have been marked in red with the corresponding gene names. The scale bar on the left indicates size in mega bases (Mb). ^a In the present study, we name the PKS gene AN8383, *ausA*, (marked in green) due to its role in the biosynthesis of austinols. ^b These ORFs have been removed from the genome annotation of the *Aspergillus* Genome Database at Stanford during preparation of Version 5. We have made the deletions according to Version 4 of annotation. ^c The annotation of this gene was taken from the genome databases at the Broad Institute of Harvard and MIT.

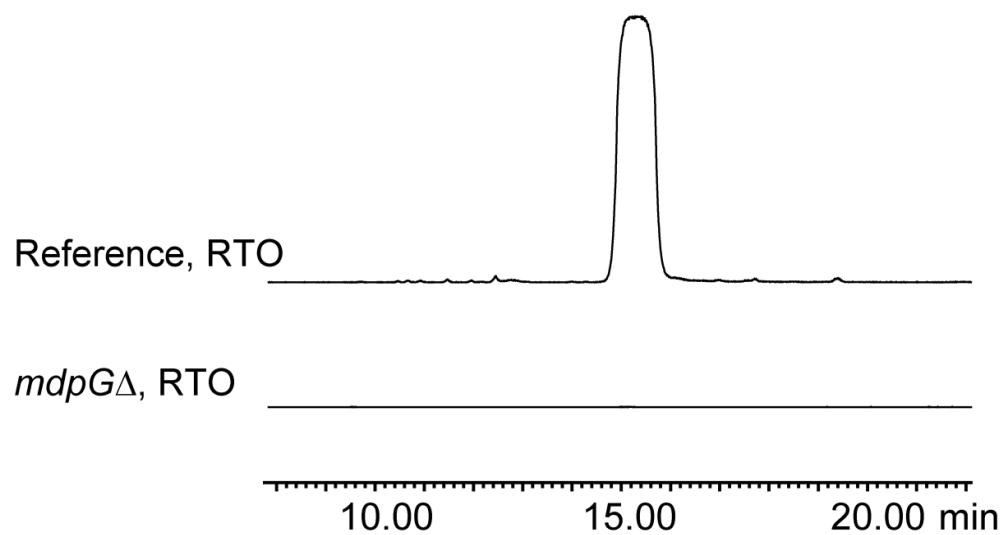
58



59

60 Fig. S5. Verification that the polyketide is absent in selected deletion mutants by either positive or
 61 negative mode extracted ion chromatograms. The *mdpG*Δ, *orsA*Δ, *easB*Δ and *stcA*Δ mutants are
 62 compared to the reference strain for the mass corresponding to the metabolites these genes are
 63 known to be involved in. (A) Monodictyphenone m/z 287 $[M-H]^-$, (B) orsellinic acid m/z 167 $[M-$
 64 $H]^-$, (C) sterigmatocystin m/z 325 $[M+H]^+$ and (D) emericellamide A m/z 610 $[M+H]^+$,
 65 emericellamide C/D m/z 596 $[M+H]^+$, emericellamide E/F m/z 624 $[M+H]^+$. * Chiang et al,
 66 (Chiang et al. 2008) also observed a minor compound with the same mass.

67

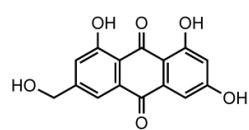


68

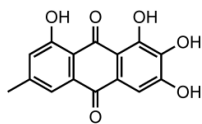
69 Fig. S6. Positive mode extracted ion chromatograms for the *mdpG*Δ strain cultivated on RTO,
70 identifying arugosin A m/z 425 $[M+H]^+$.

71

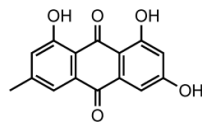
72



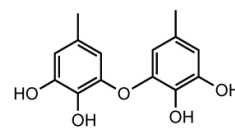
ω-hydroxyemodin (7)



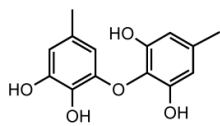
2-hydroxymodin (8)



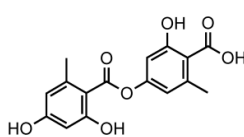
Emodin (9)



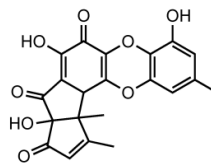
Violaceol I (13)



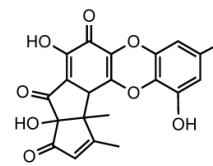
Violaceol II (14)



Lecanoric acid (15)



F-9775A (16)



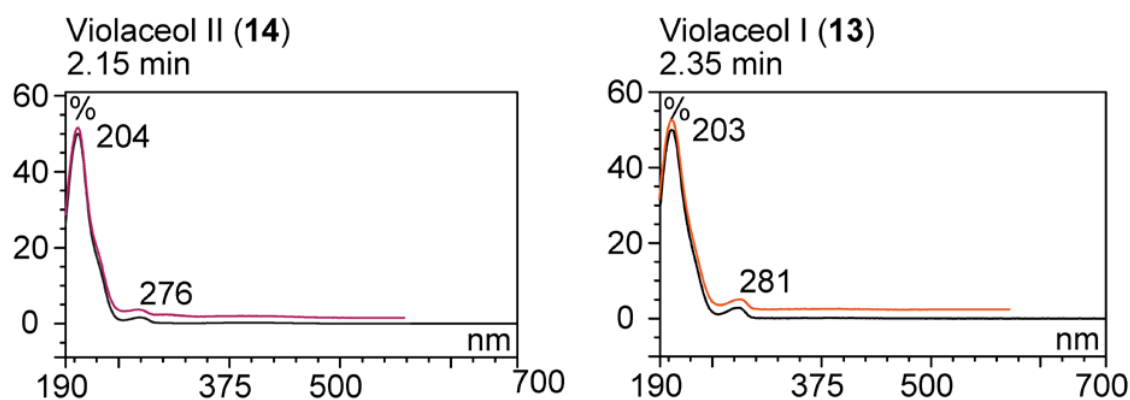
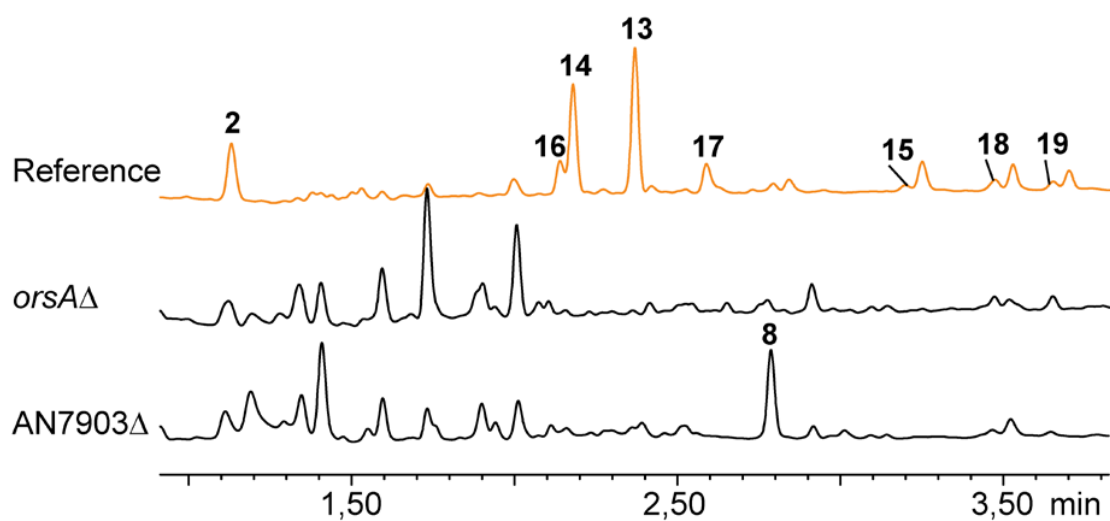
F-9775B (17)

73

74 Fig. S7. Additional polyketides that were detected in metabolite extracts in this study.

75

76



77

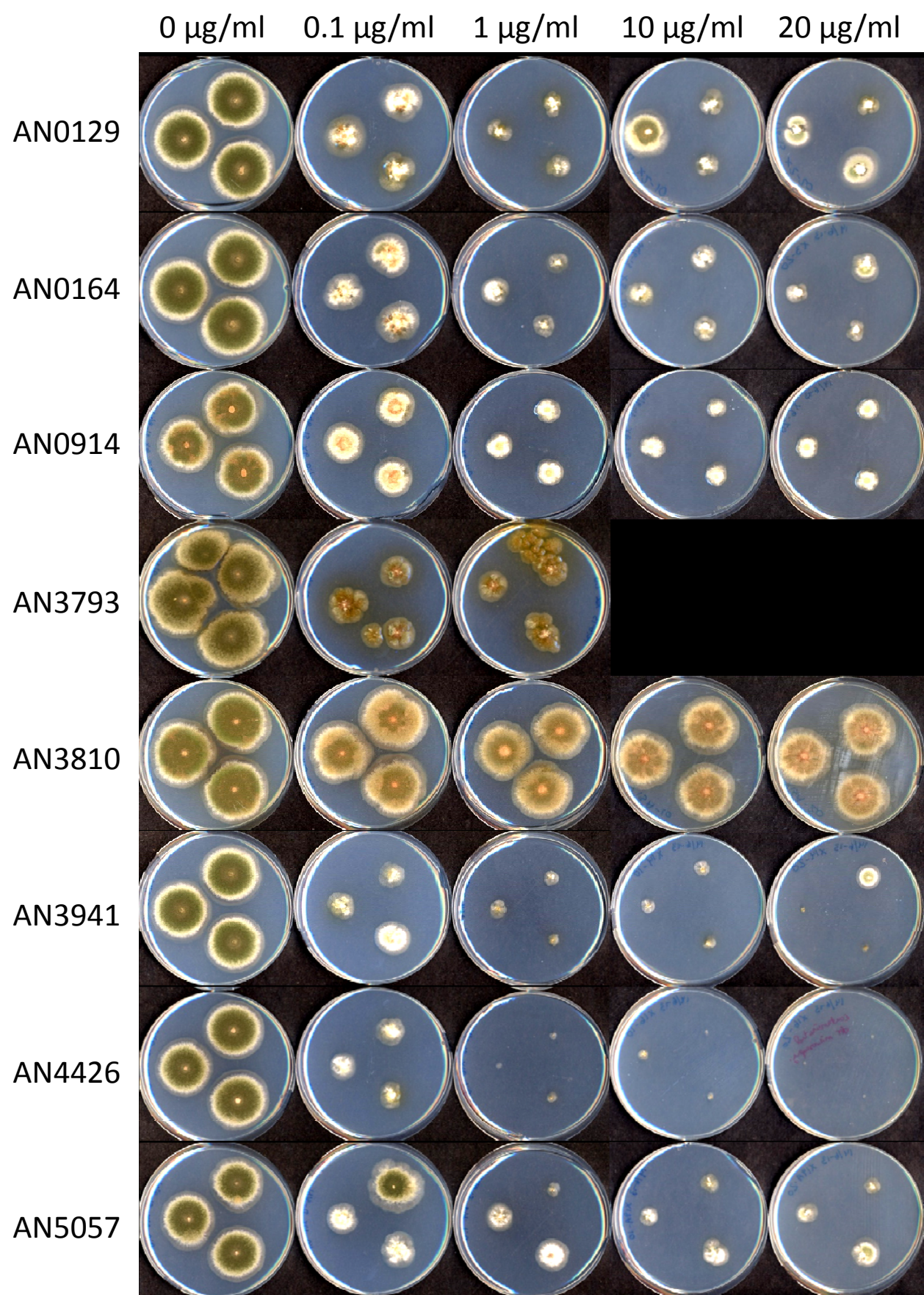
78 Fig. S8. (A) UHPLC profiles (UV 210 nm) of *orsAΔ*, AN7903Δ compared to the reference strain.
 79 The peaks are numbered as follows: orsellinic acid (**2**), ω-hydroxyemodin (**8**), violaceol I (**13**),
 80 violaceol II (**14**), lecanoric acid (**15**), F9775A (**16**), F9775B (**17**), austinol (**18**), dehydroaustinol
 81 (**19**). (B) UV-spectra of violaceol I (left panel) and II (right panel) standard are shown in black, with
 82 an overlay shown in red of the UV-spectra of the sample representing peak 14 (left panel) and peak
 83 13 (right panel). There is agreement between standards and samples.

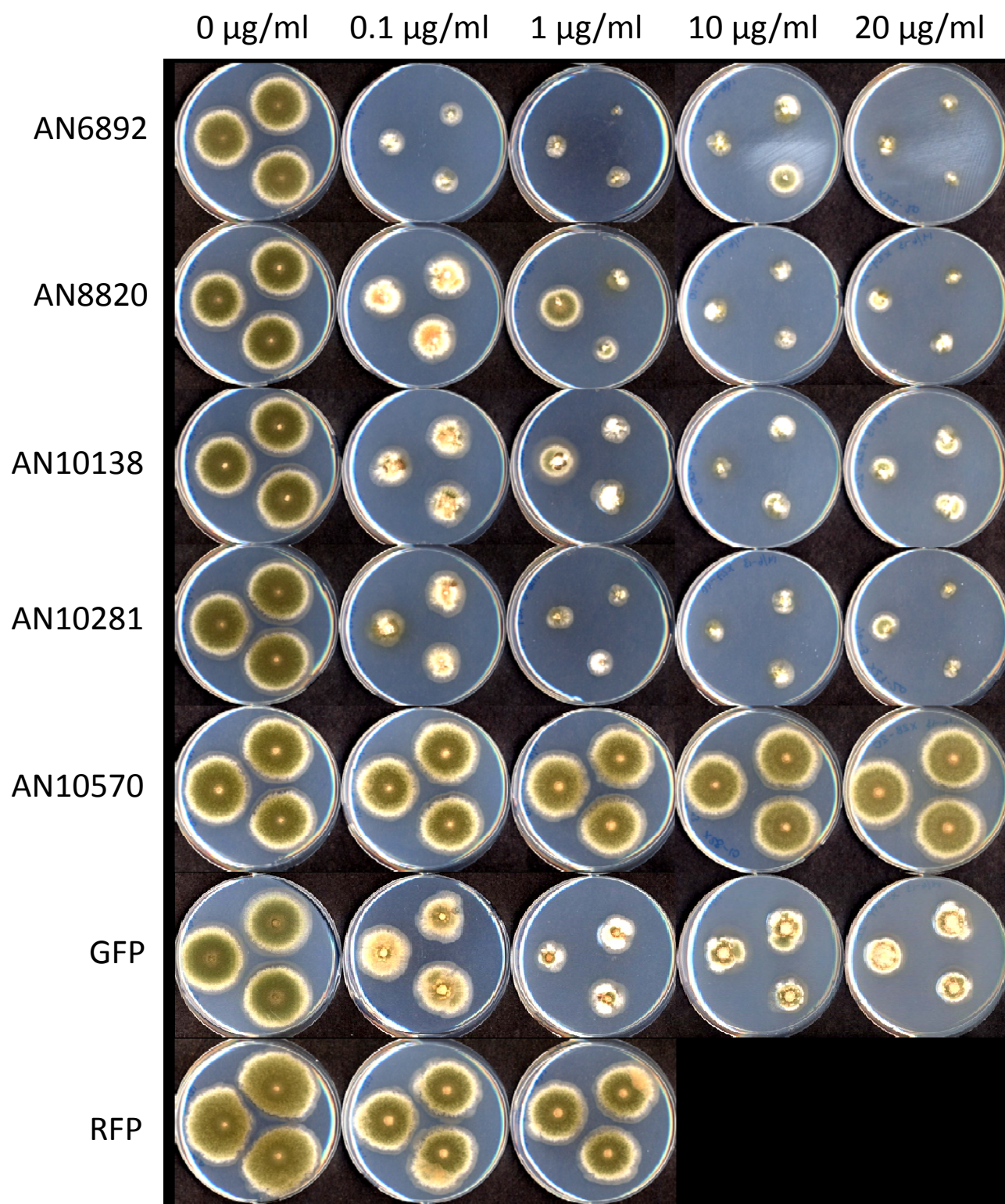
84

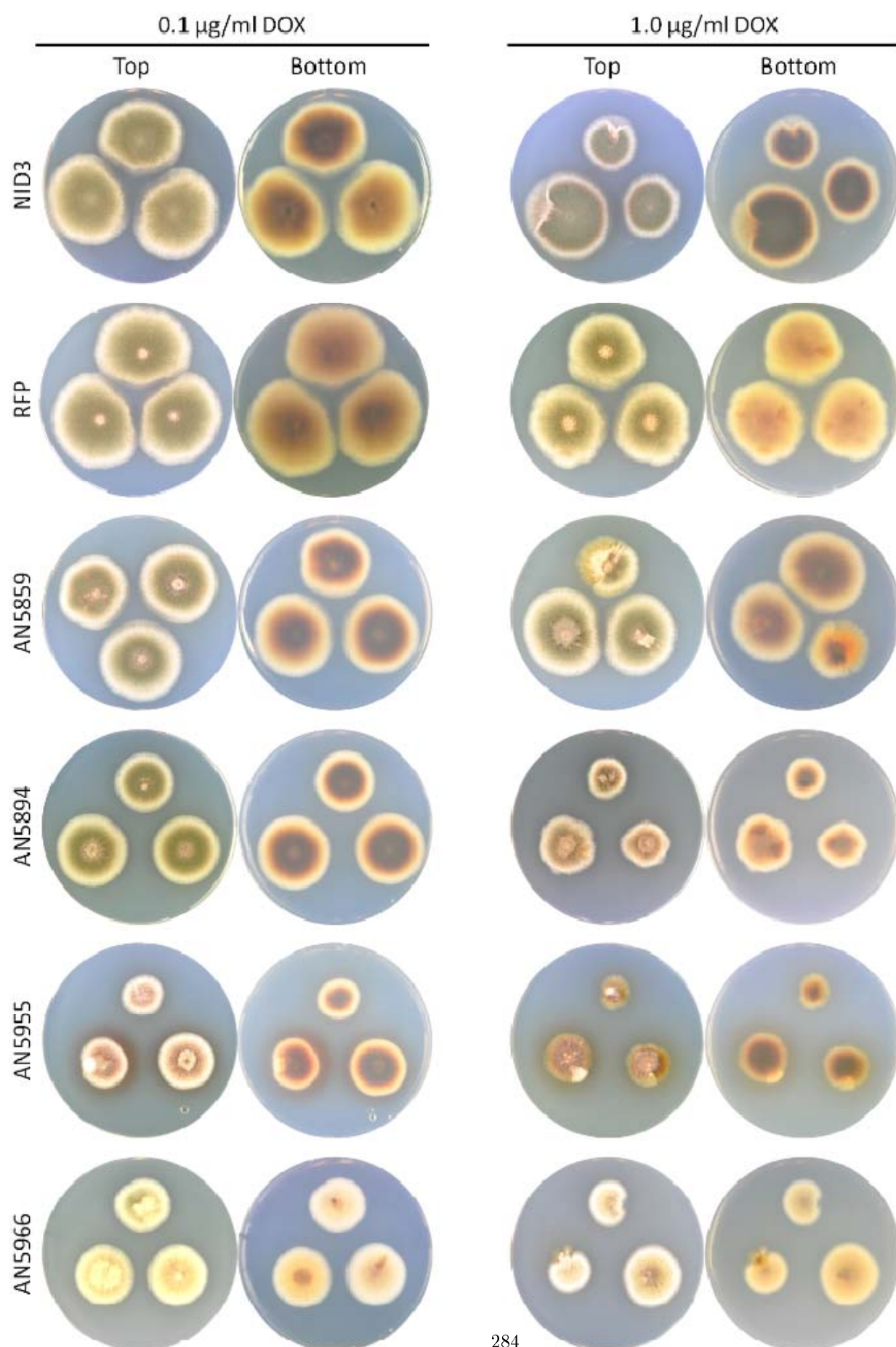
Appendix B

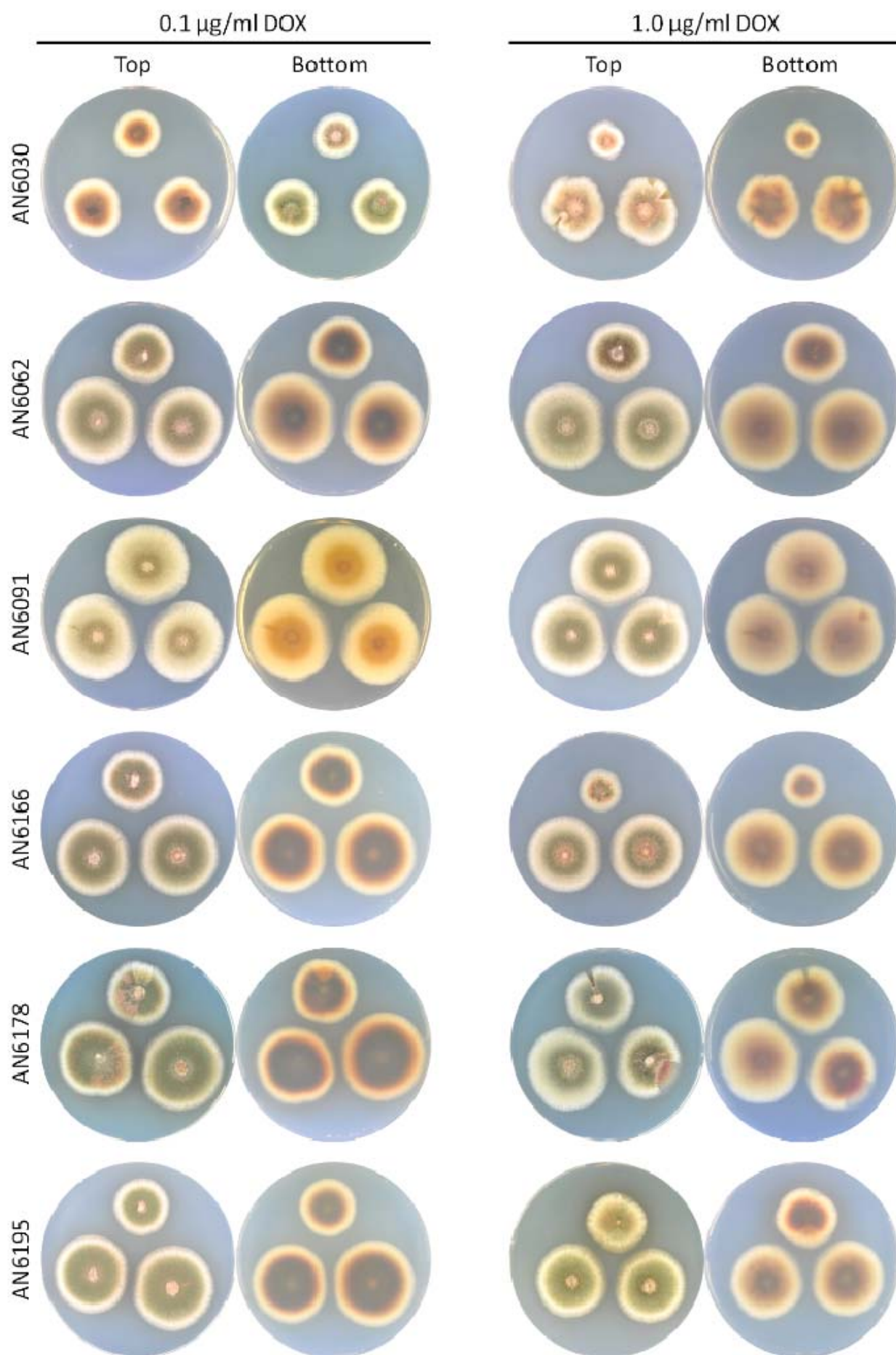
Supplementaries for Chapter 6

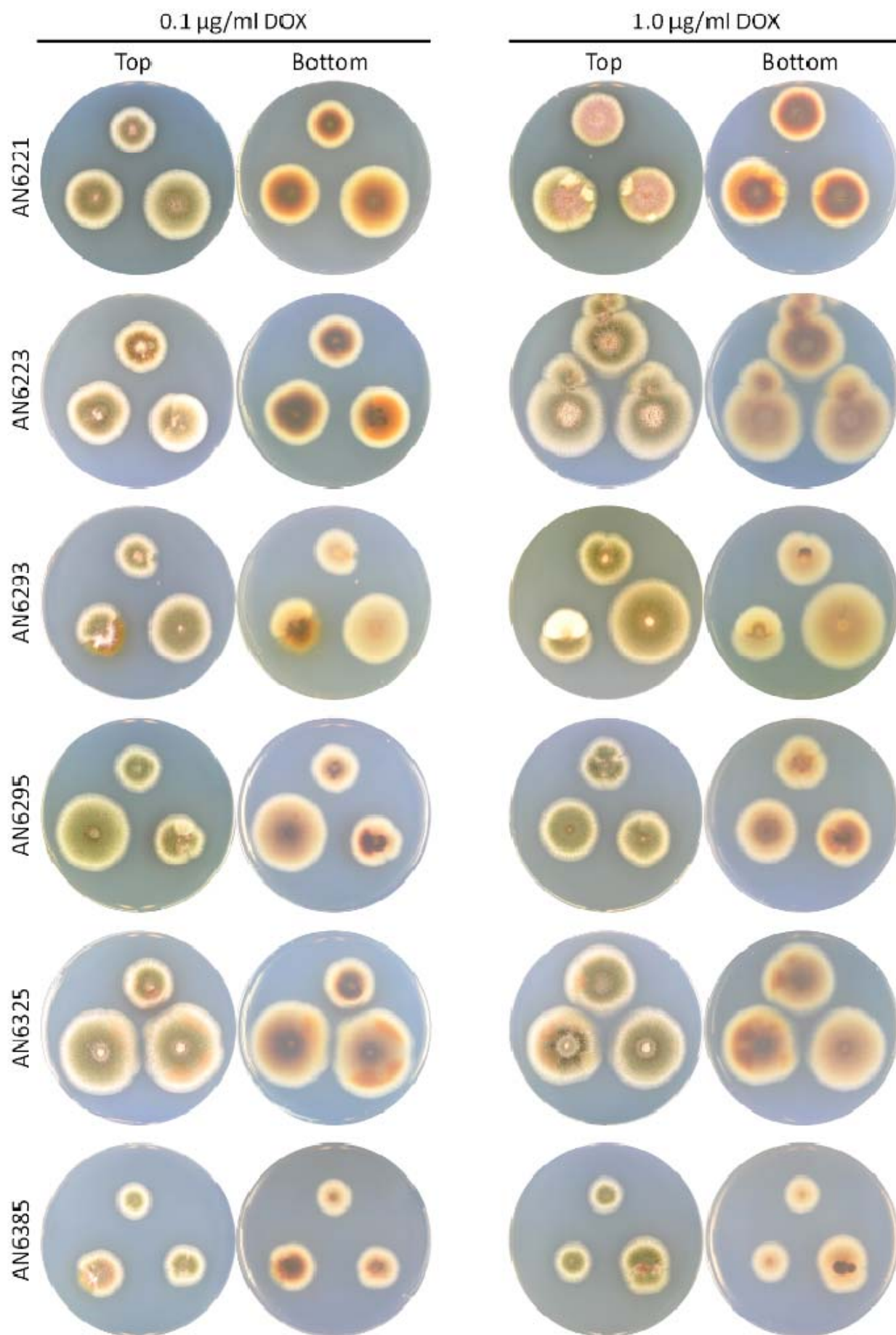
B.1 Large pictures of phosphatase and TF strains

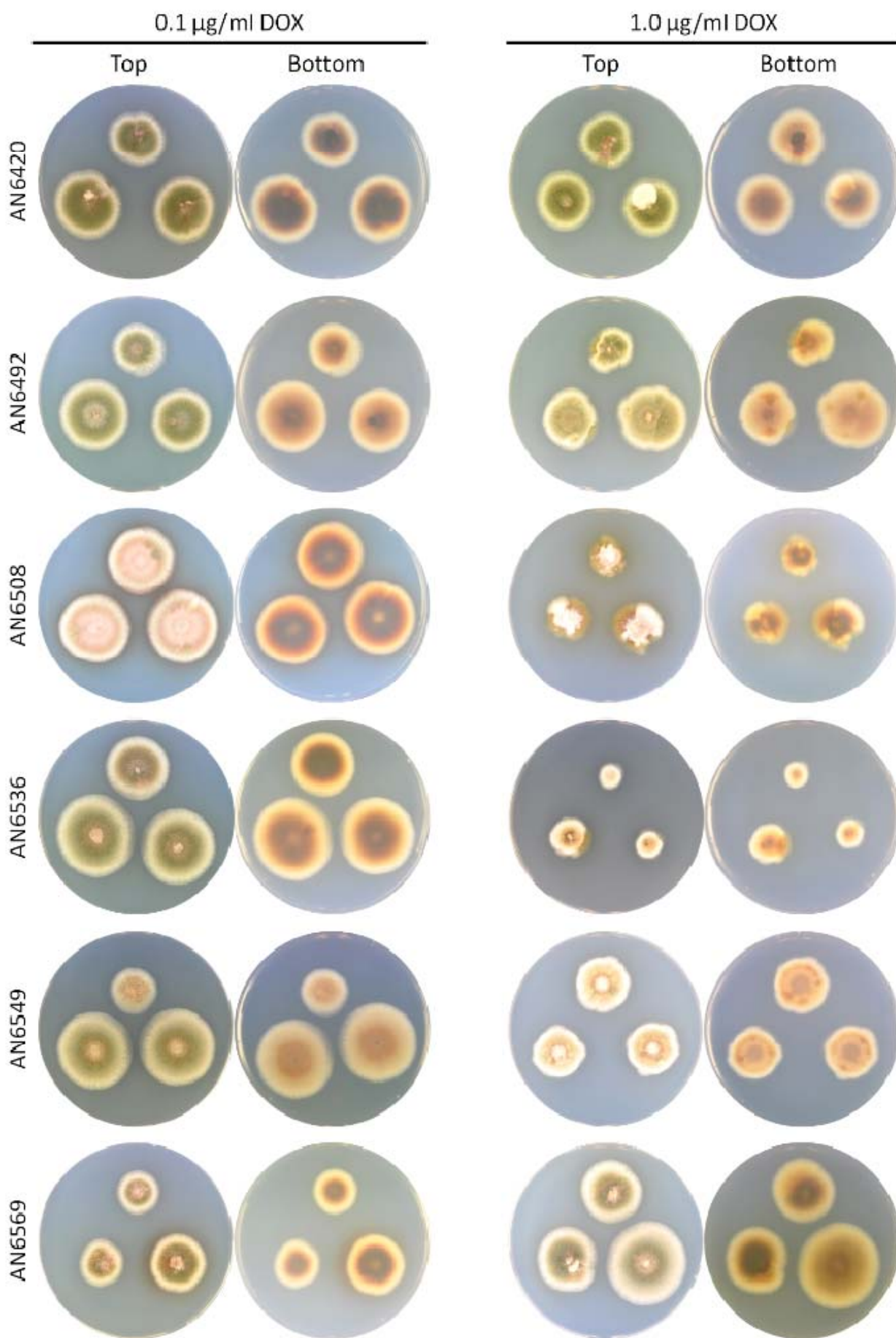


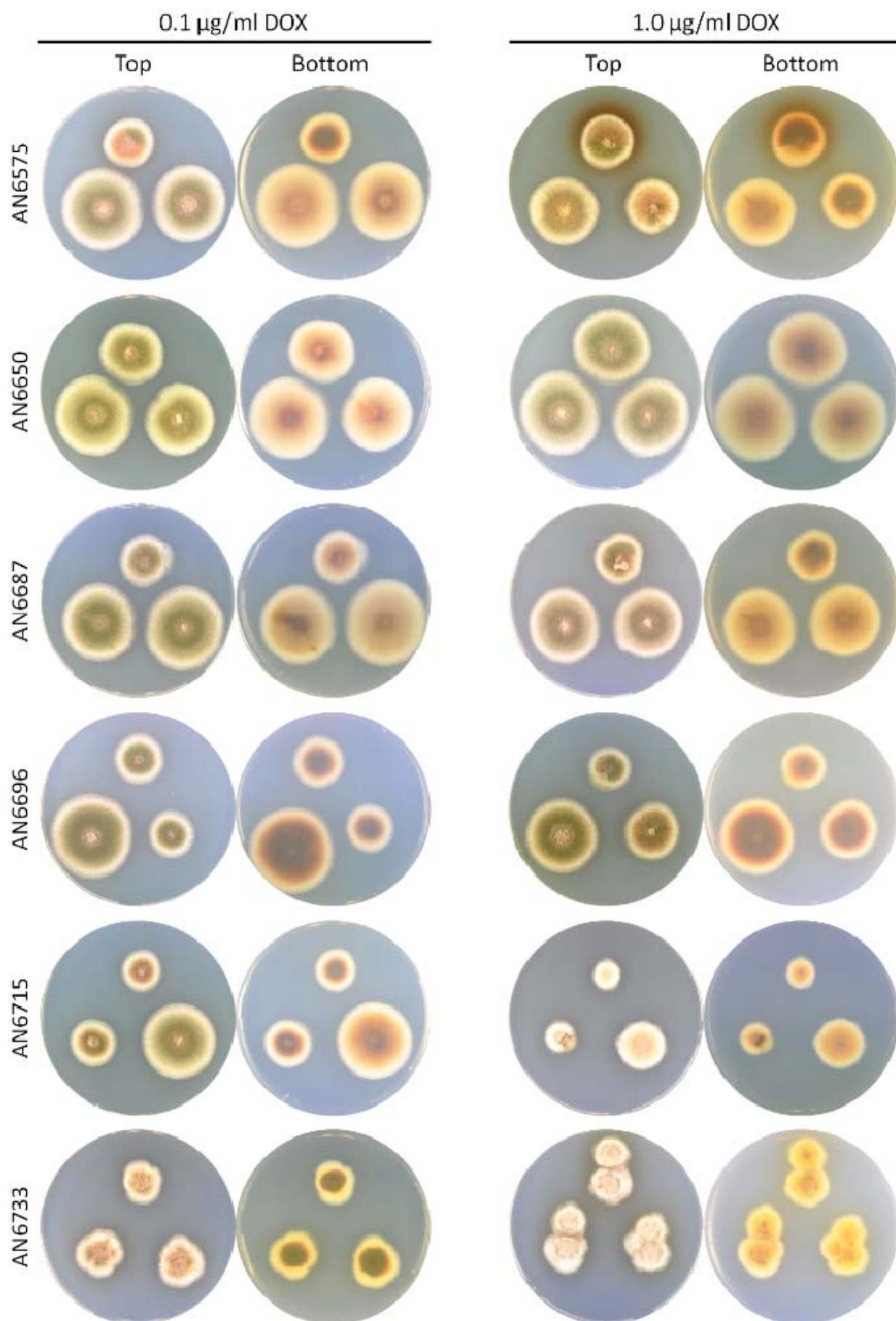


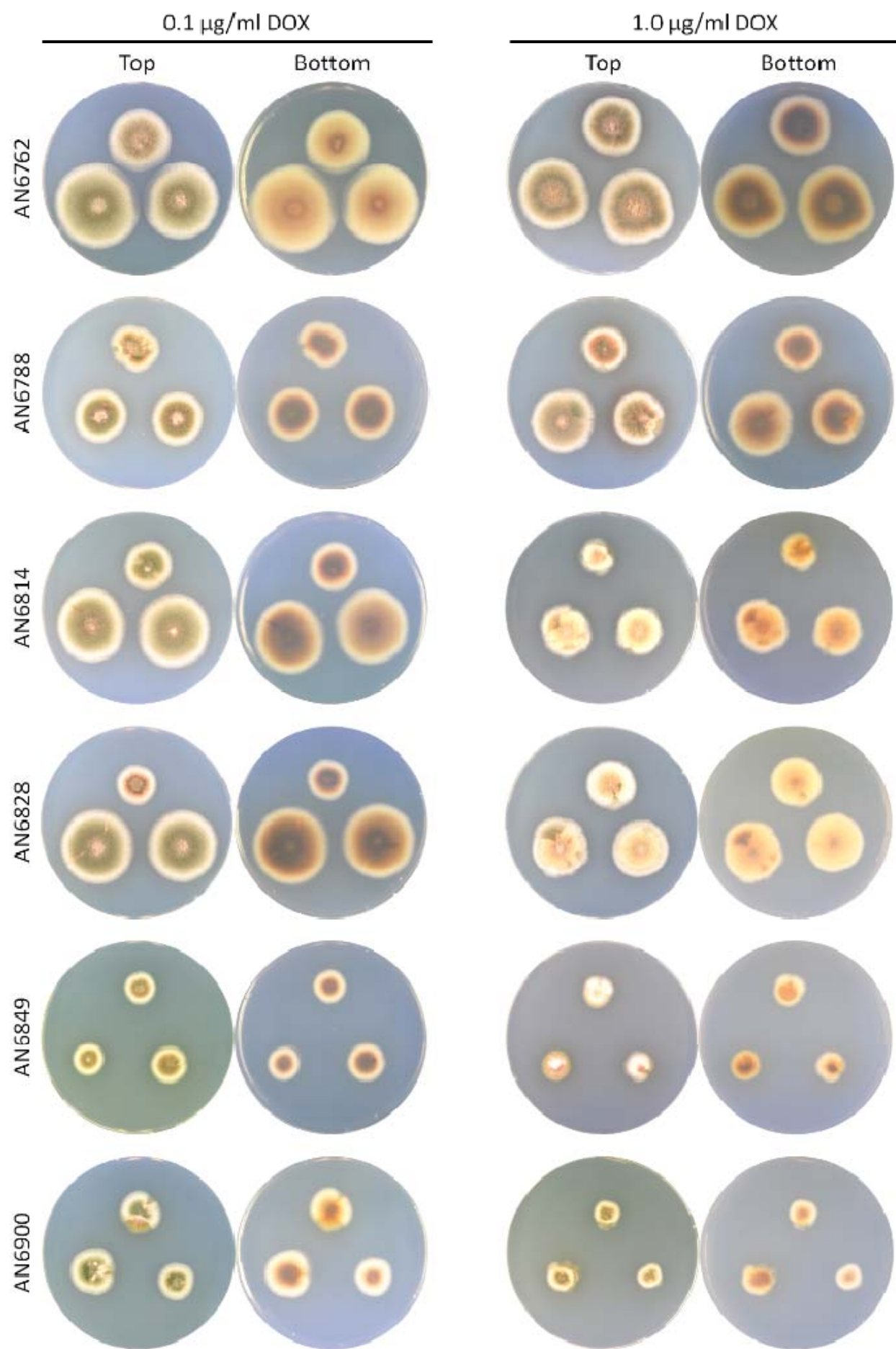


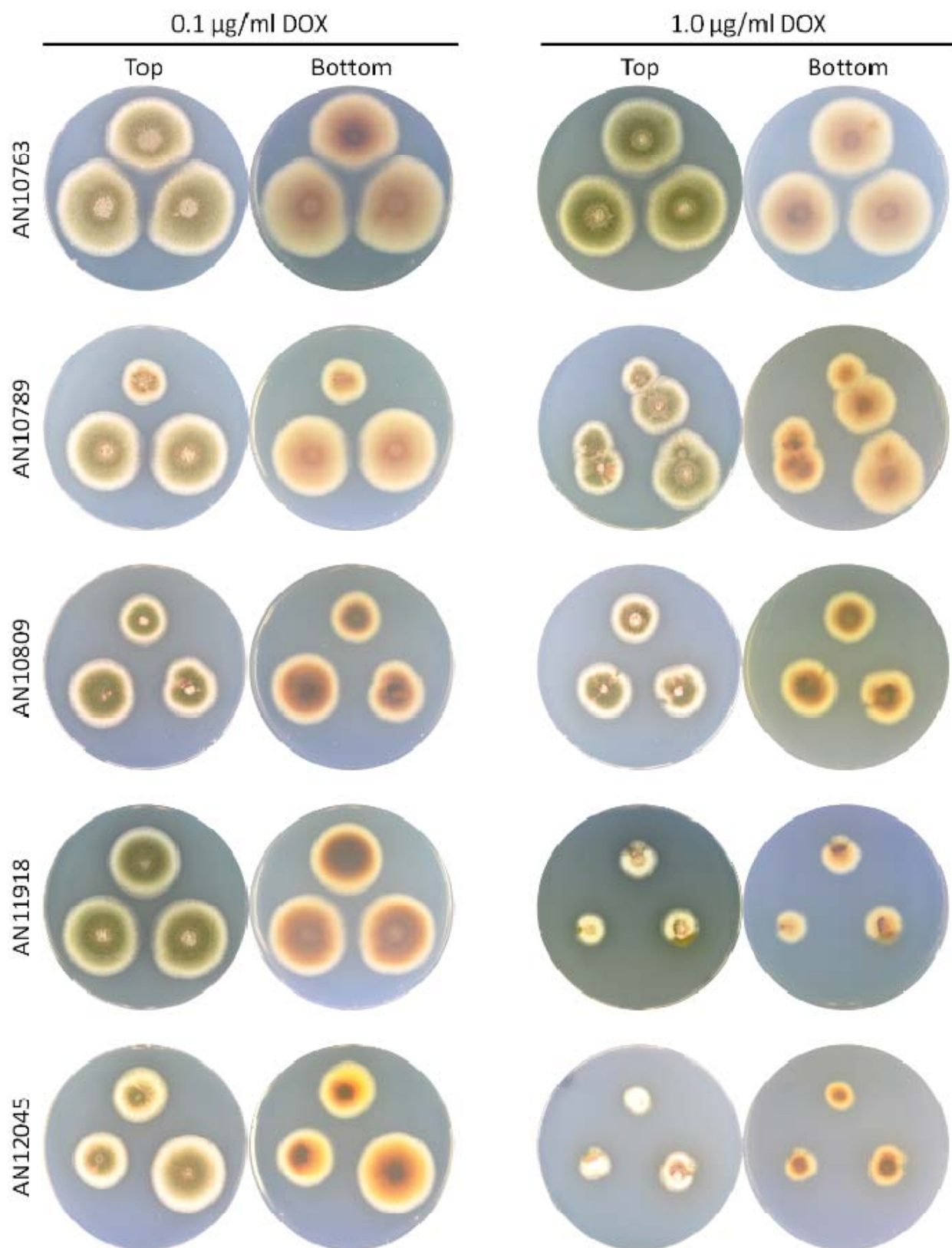










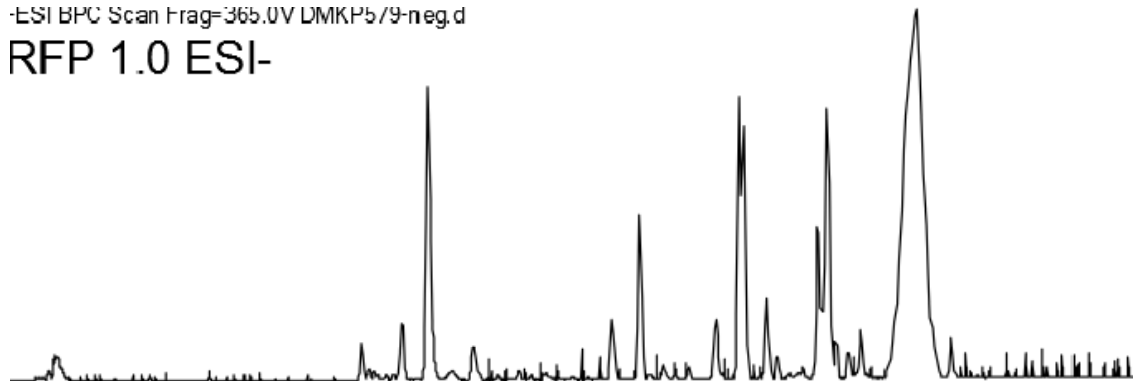


B.2 Chromatograms of phosphatase and TF strains

This section contains ESI-, ESI+, and DAD chromatograms of phosphatase and TF strains, respectively. Black chromatograms are ESI- base peak chromatograms (BPC) for all ions, in a m/z 100-1700 range. Blue chromatograms are ESI+ base peak chromatograms (BPC) for all ions except $m/z = 186.2 \pm 0.05$, in a m/z 100-1000 range. The ion corresponding to $m/z = 186.2$ represents the ion pair reagent tetrapropylammonium (TPA), which was used in a sample run previous to these analyses. DAD chromatograms (measured at 280 nm) are shown in green.

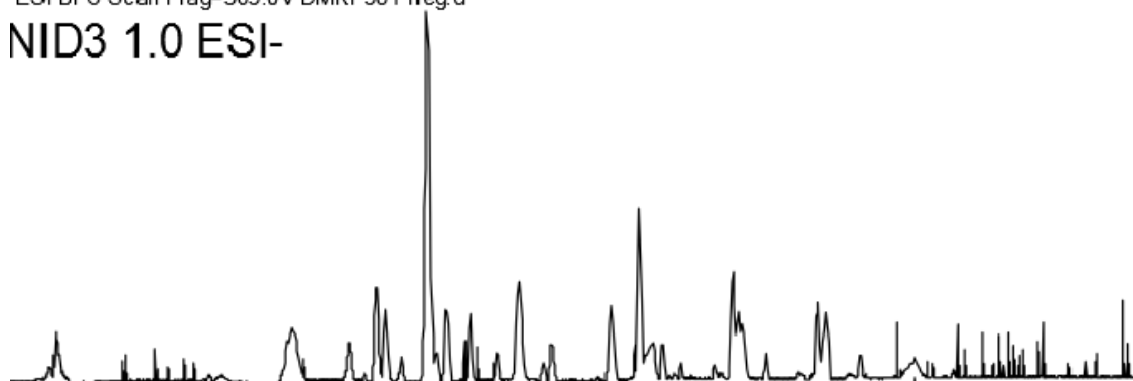
-ESI BPC Scan Frag=365.0V DMKP5/9-neg.d

RFP 1.0 ESI-



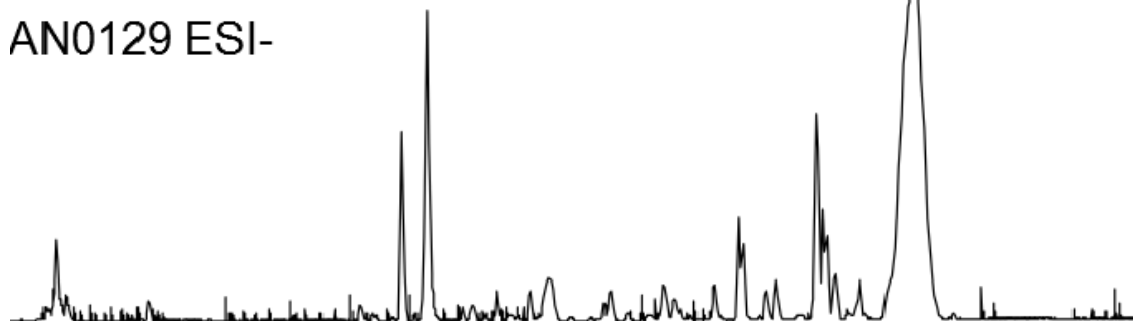
-ESI BPC Scan Frag=365.0V DMKP581-neg.d

NID3 1.0 ESI-



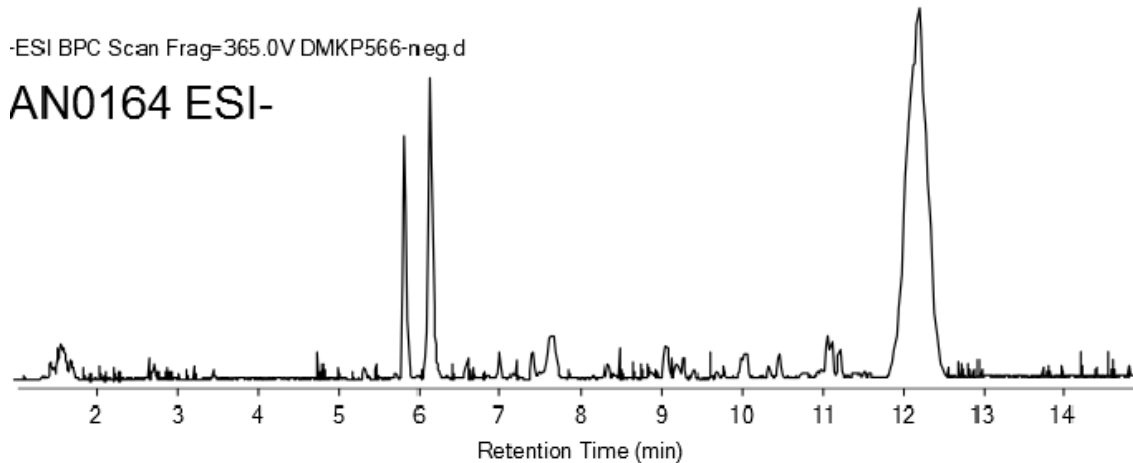
-ESI BPC Scan Frag=365.0V DMKP565-neg.d

AN0129 ESI-



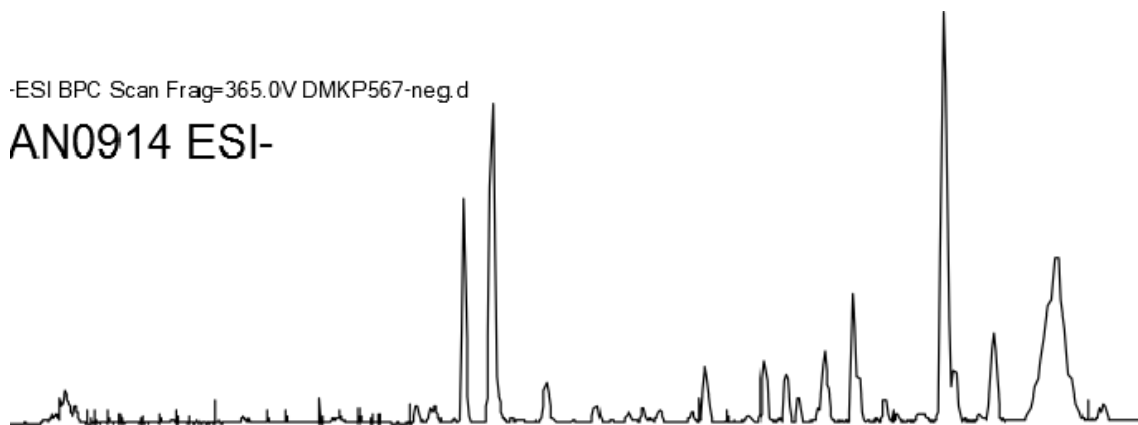
-ESI BPC Scan Frag=365.0V DMKP566-neg.d

AN0164 ESI-



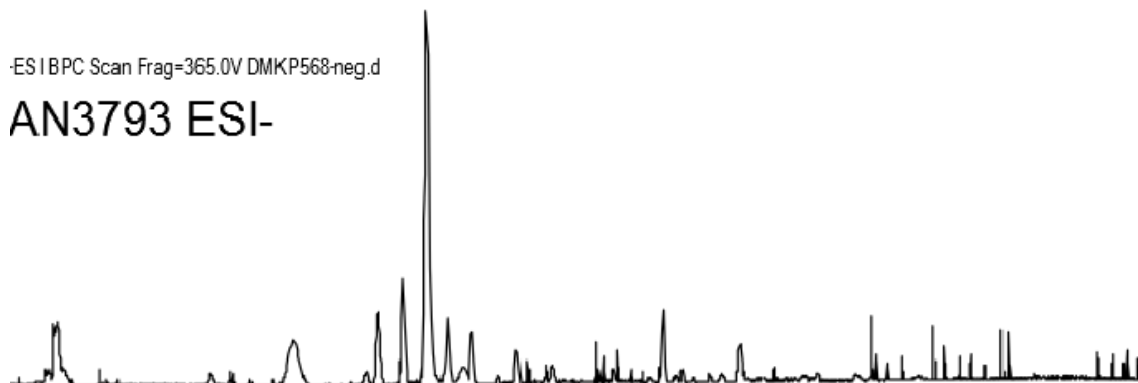
-ESI BPC Scan Frag=365.0V DMKP567-neg.d

AN0914 ESI-



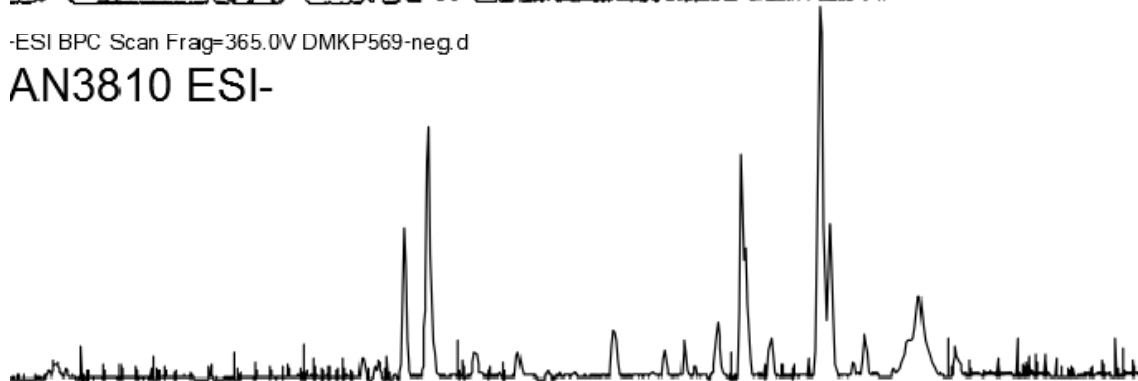
-ESI BPC Scan Frag=365.0V DMKP568-neg.d

AN3793 ESI-



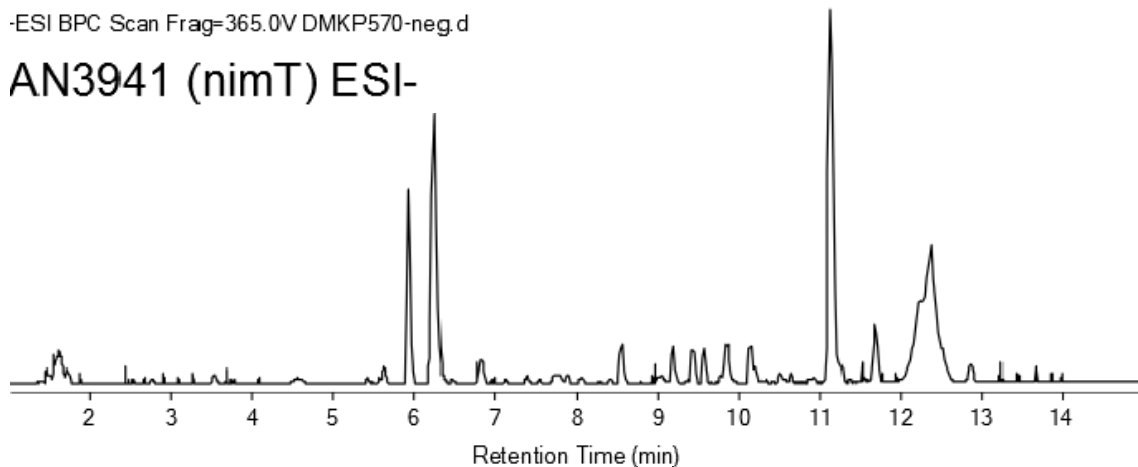
-ESI BPC Scan Frag=365.0V DMKP569-neg.d

AN3810 ESI-



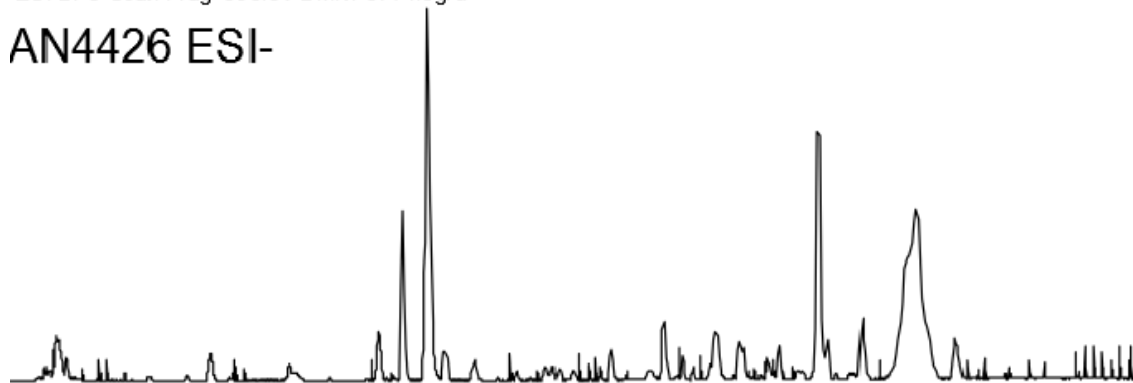
-ESI BPC Scan Frag=365.0V DMKP570-neg.d

AN3941 (nimT) ESI-



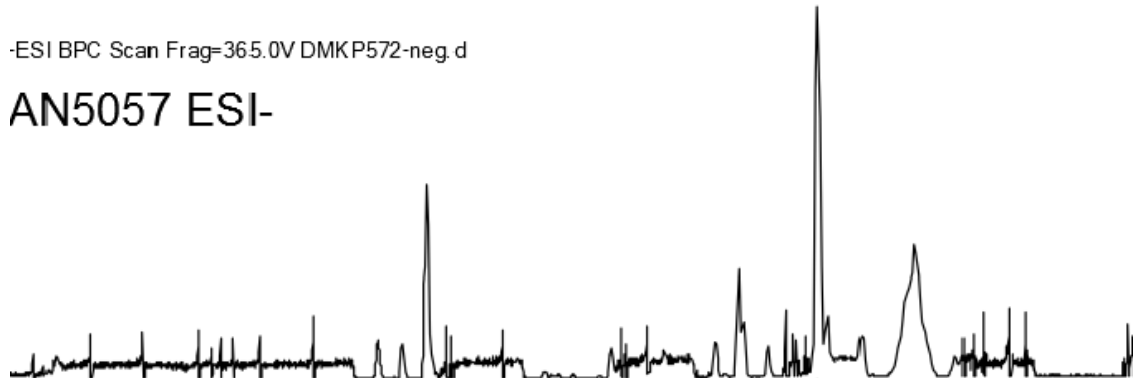
-ESI BPC Scan Frag=365.0V DMKP571-neg.d

AN4426 ESI-



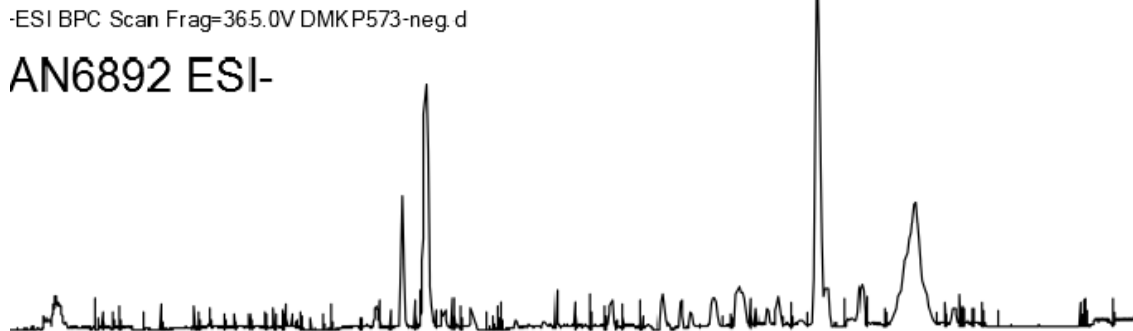
-ESI BPC Scan Frag=365.0V DMKP572-neg.d

AN5057 ESI-



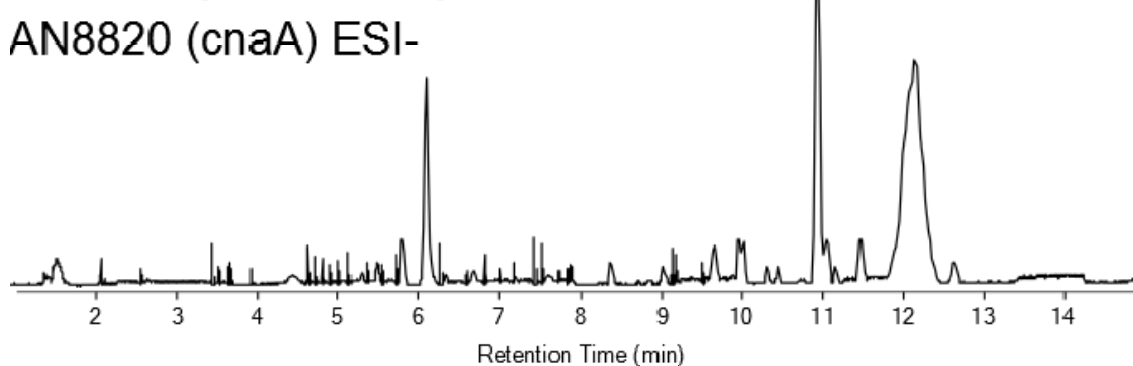
-ESI BPC Scan Frag=365.0V DMKP573-neg.d

AN6892 ESI-



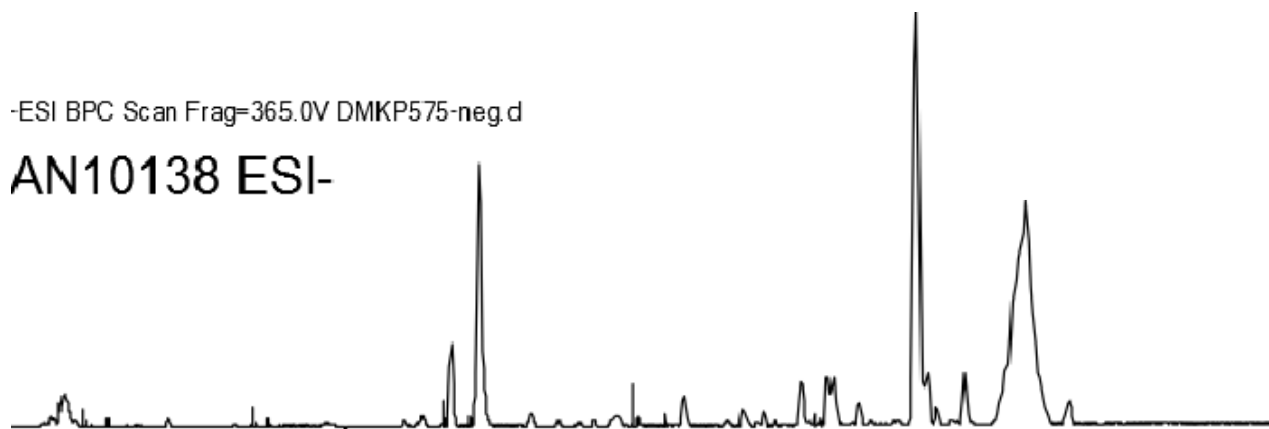
-ESI BPC Scan Frag=365.0V DMKP574-neg.d

AN8820 (cnaA) ESI-



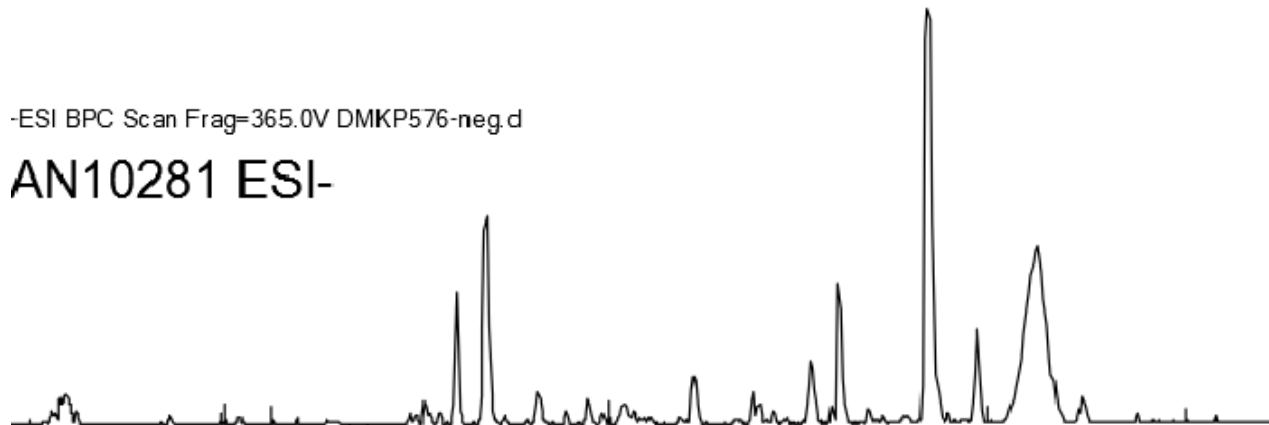
-ESI BPC Scan Frag=365.0V DMKP575-neg.d

AN10138 ESI-



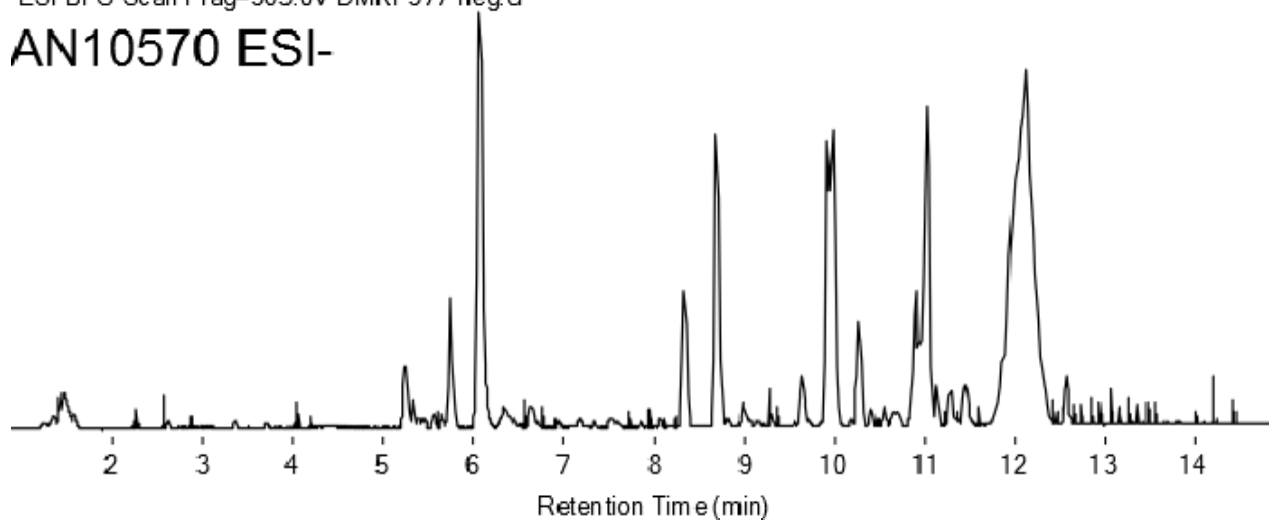
-ESI BPC Scan Frag=365.0V DMKP576-neg.d

AN10281 ESI-



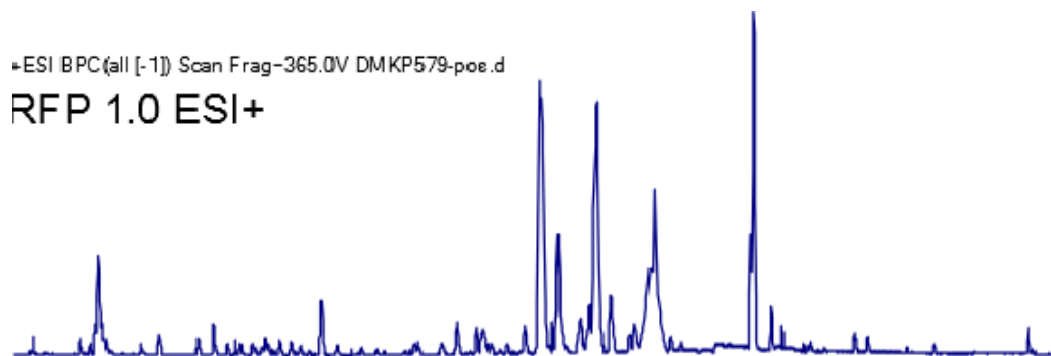
-ESI BPC Scan Frag=365.0V DMKP577-neg.d

AN10570 ESI-



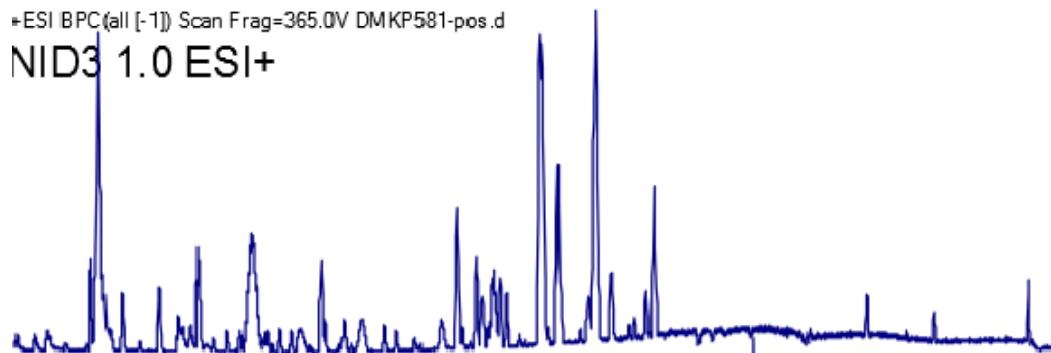
*ESI BPC(all [-1]) Scan Frag=365.0V DMKP579-pos.d

RFP 1.0 ESI+



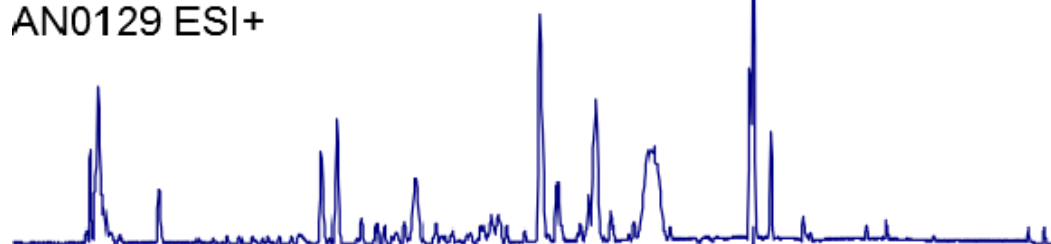
*ESI BPC(all [-1]) Scan Frag=365.0V DMKP581-pos.d

NID3 1.0 ESI+



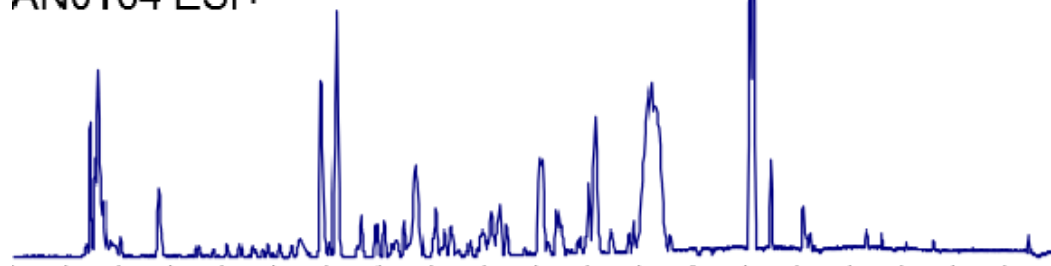
*ESI BPC(all [-1]) Scan Frag=365.0V DMKP565-pos.d

AN0129 ESI+



*ESI BPC(all [-1]) Scan Frag=365.0V DMKP566-pos.d

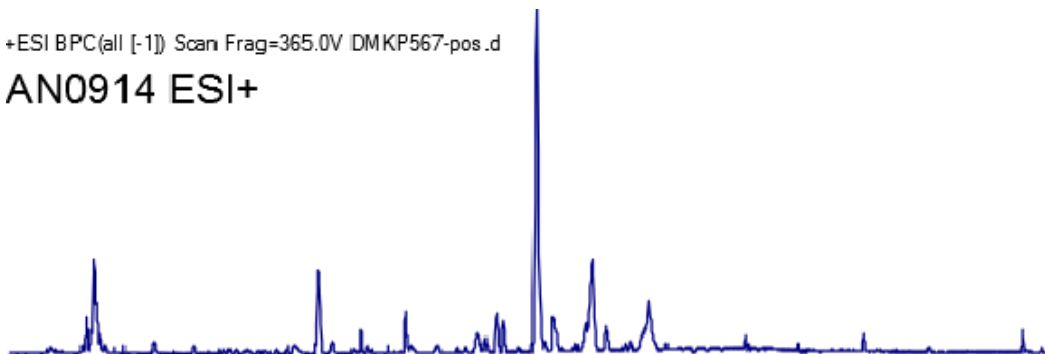
AN0164 ESI+



Retention time (min)

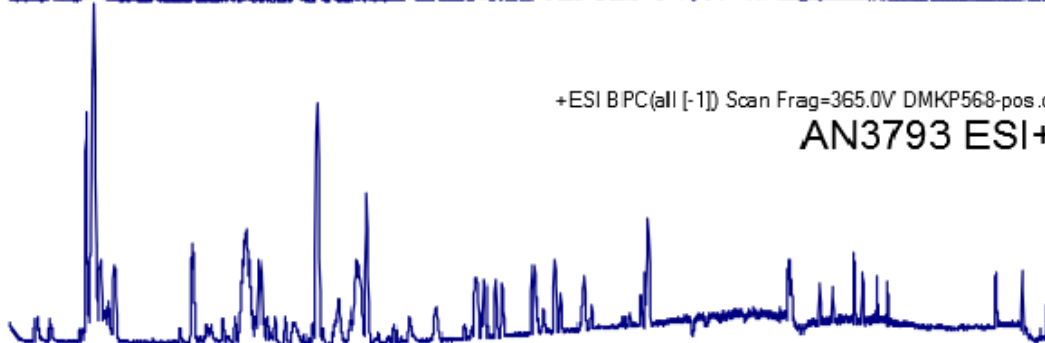
+ESI BPC(all [-1]) Scan Frag=365.0V DMKP567-pos.d

AN0914 ESI+



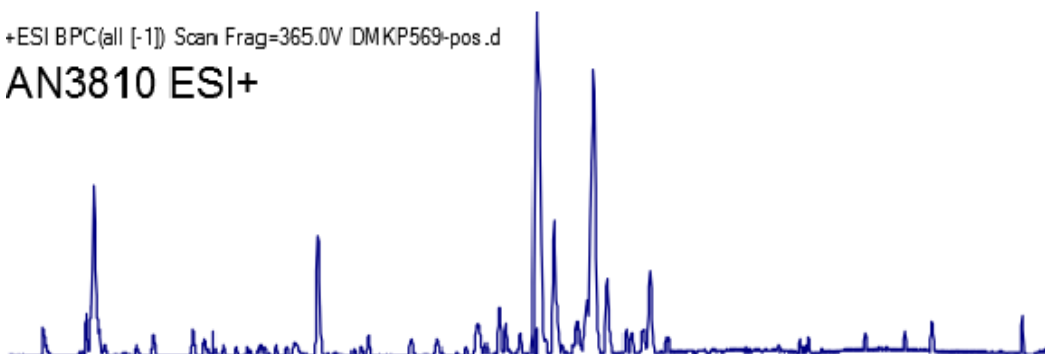
+ESI BPC(all [-1]) Scan Frag=365.0V DMKP568-pos.d

AN3793 ESI+



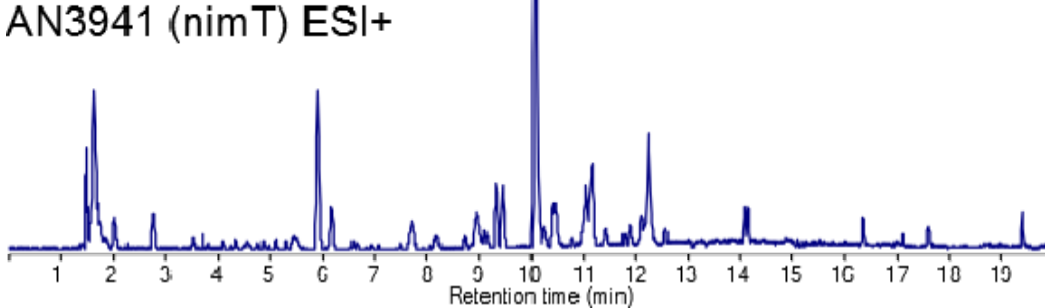
+ESI BPC(all [-1]) Scan Frag=365.0V DMKP569-pos.d

AN3810 ESI+



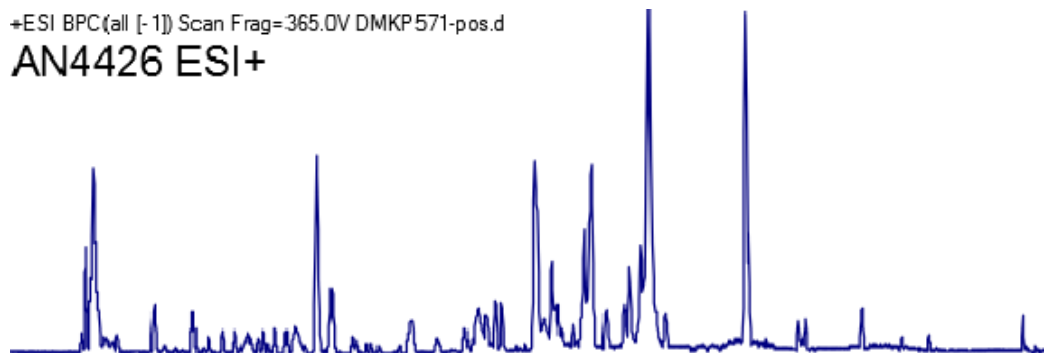
+ESI BPC(all [-1]) Scan Frag=365.0V DMKP570-pos.d

AN3941 (nimT) ESI+



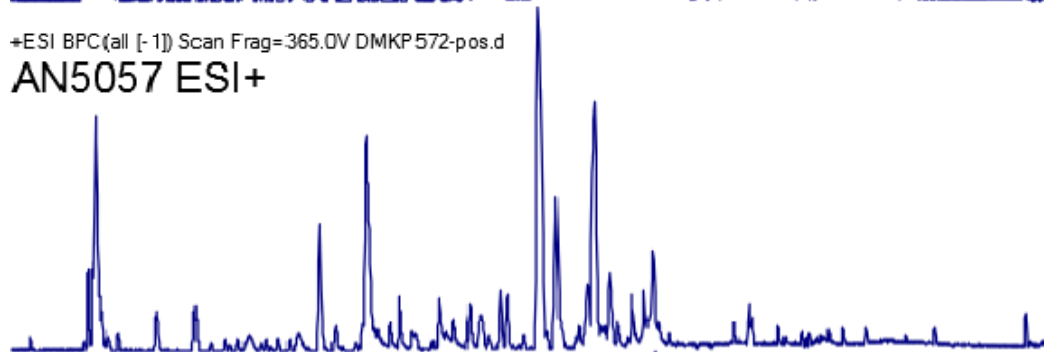
+ESI BPC(all [-1]) Scan Frag=365.0V DMKP571-pos.d

AN4426 ESI+



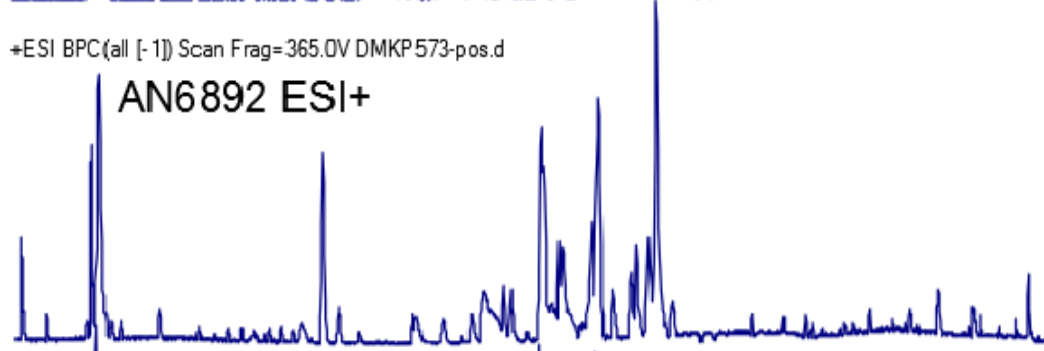
+ESI BPC(all [-1]) Scan Frag=365.0V DMKP572-pos.d

AN5057 ESI+



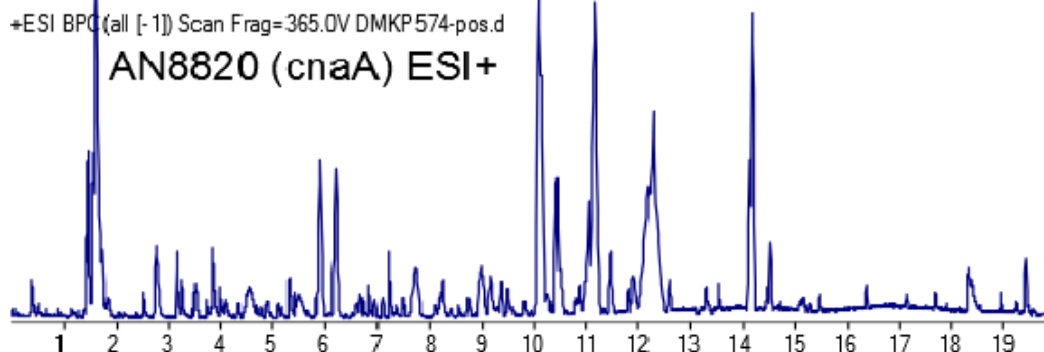
+ESI BPC(all [-1]) Scan Frag=365.0V DMKP573-pos.d

AN6892 ESI+



+ESI BPC(all [-1]) Scan Frag=365.0V DMKP574-pos.d

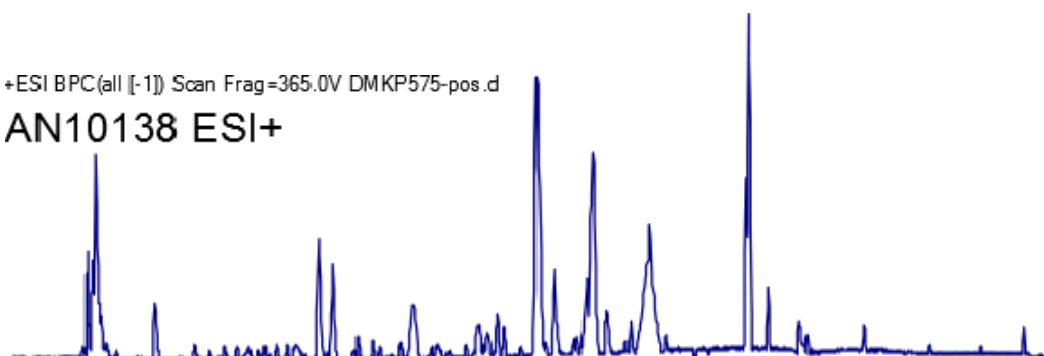
AN8820 (cnaA) ESI+



Retention time (min)

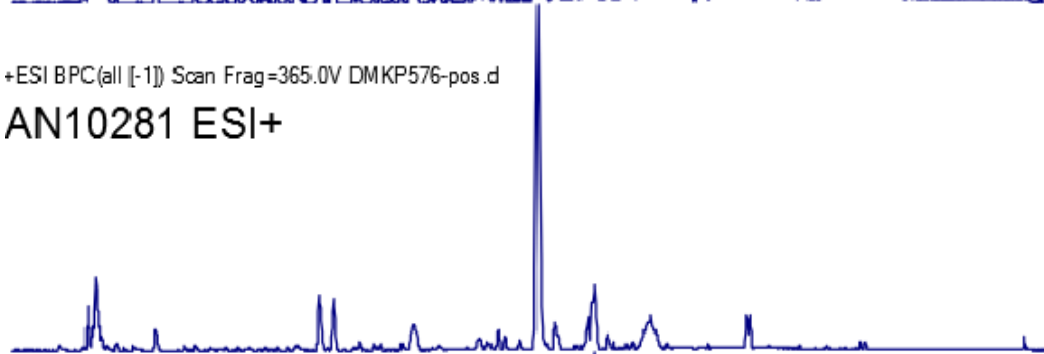
+ESI BPC(all [-1]) Scan Frag=365.0V DMKP575-pos.d

AN10138 ESI+



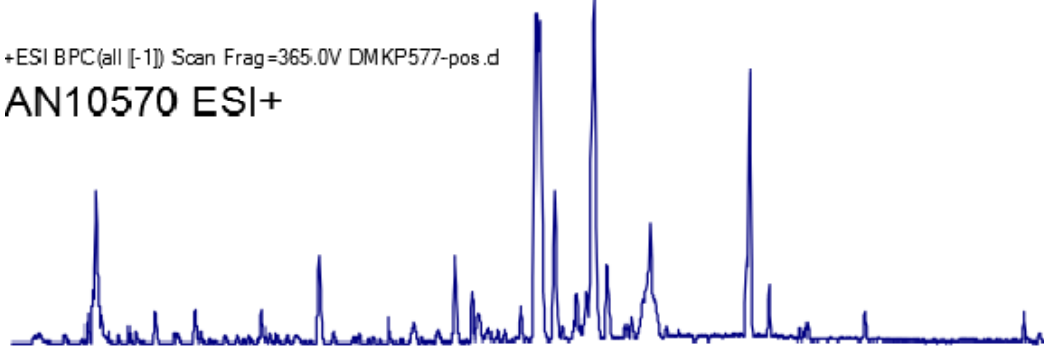
+ESI BPC(all [-1]) Scan Frag=365.0V DMKP576-pos.d

AN10281 ESI+



+ESI BPC(all [-1]) Scan Frag=365.0V DMKP577-pos.d

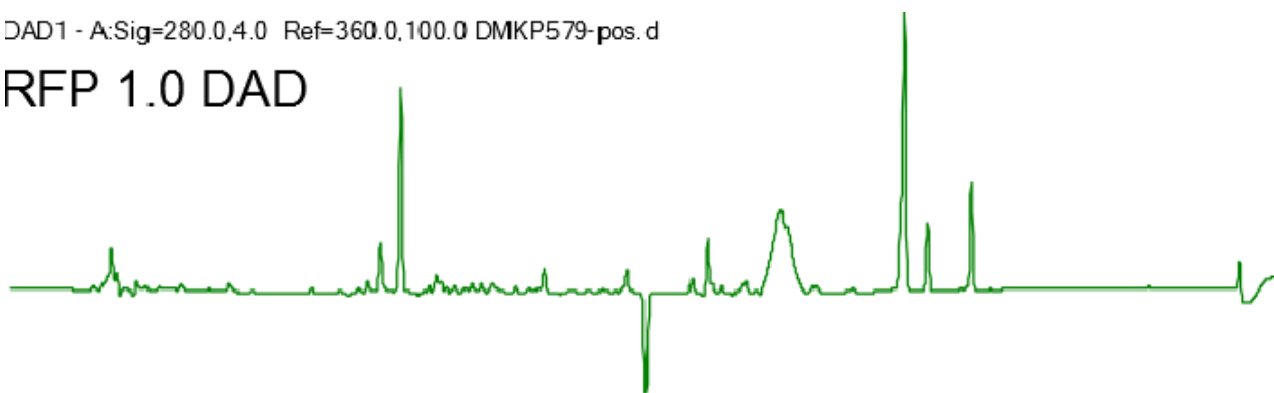
AN10570 ESI+



Retention time (min)

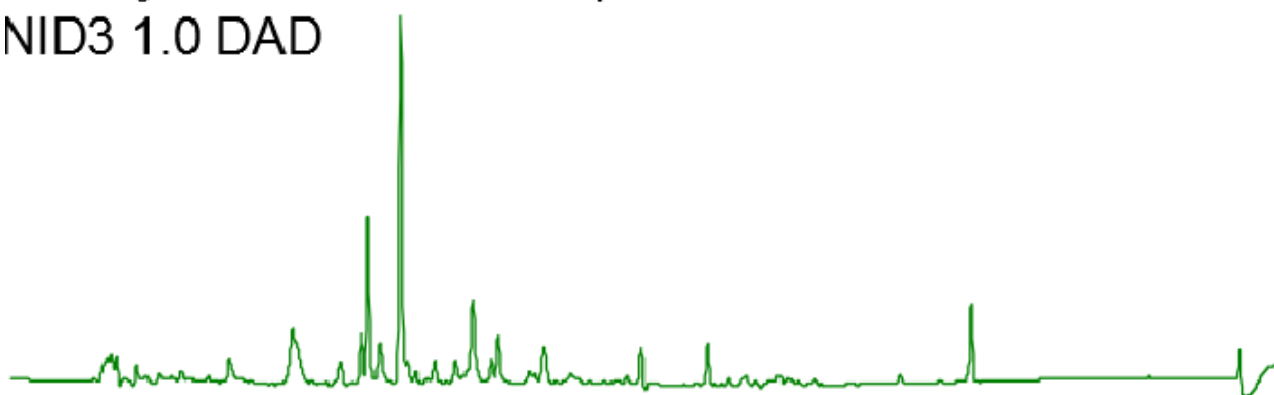
DAD1 - A:Sig=280.0,4.0 Ref=360.0,100.0 DMKP579-pos. d

RFP 1.0 DAD



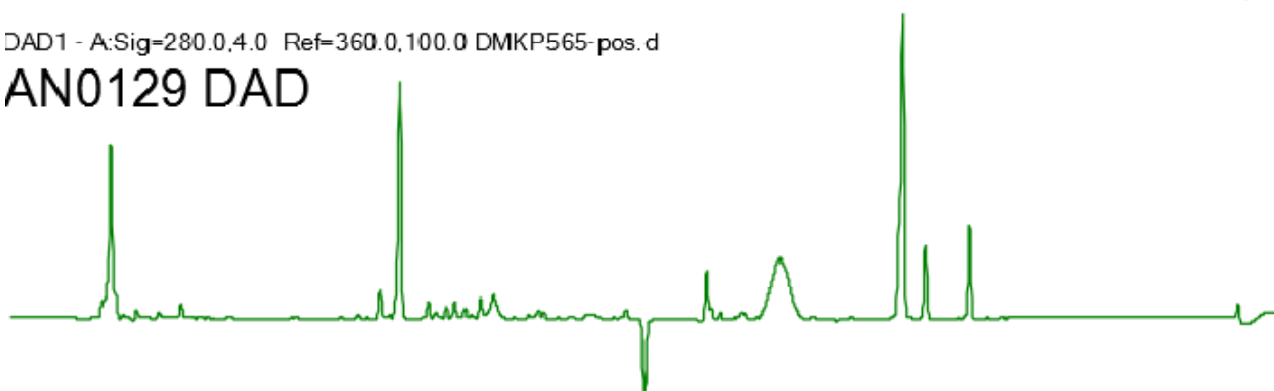
DAD1 - A:Sig=280.0,4.0 Ref=360.0,100.0 DMKP581-pos. d

NID3 1.0 DAD



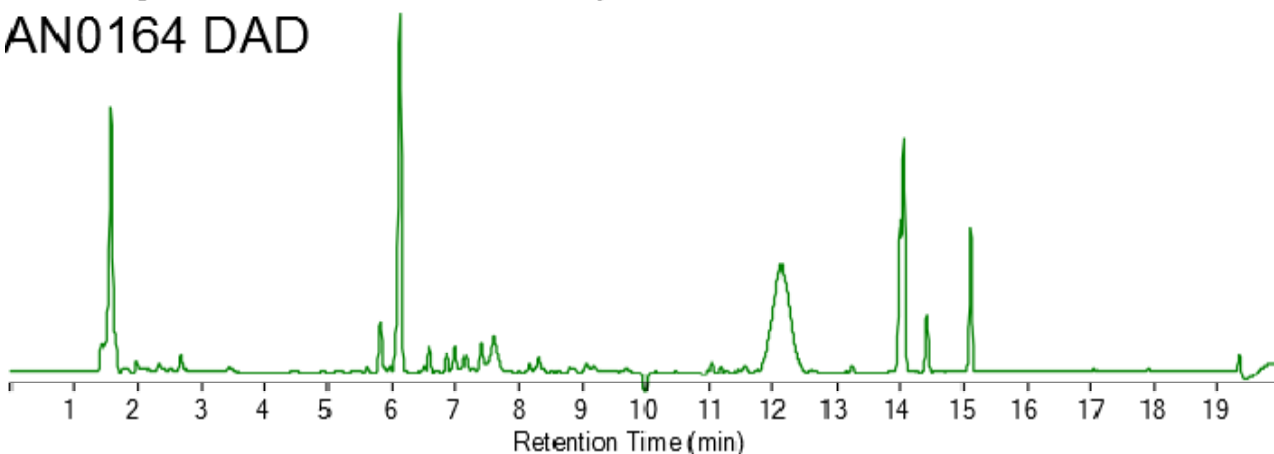
DAD1 - A:Sig=280.0,4.0 Ref=360.0,100.0 DMKP565-pos. d

AN0129 DAD



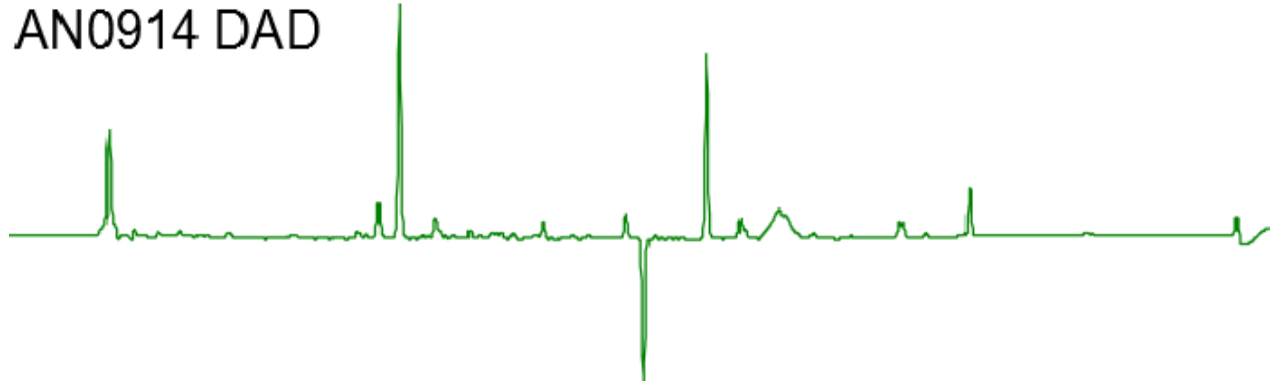
DAD1 - A:Sig=280.0,4.0 Ref=360.0,100.0 DMKP566-pos. d

AN0164 DAD



DAD1 - A:Sig=280.0,4.0 Ref=360.0,100.0 DMKP567-pos.d

AN0914 DAD



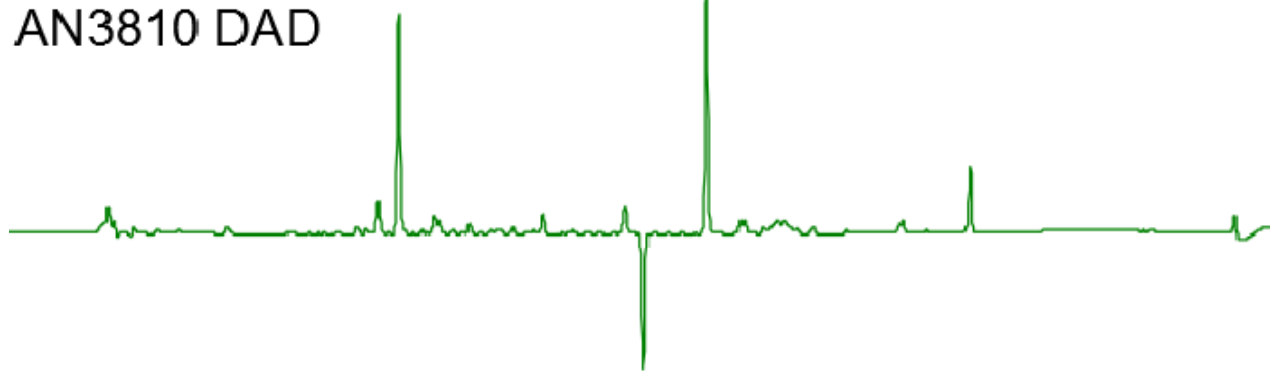
DAD1 - A:Sig=280.0,4.0 Ref=360.0,100.0 DMKP568-pos.d

AN3793 DAD



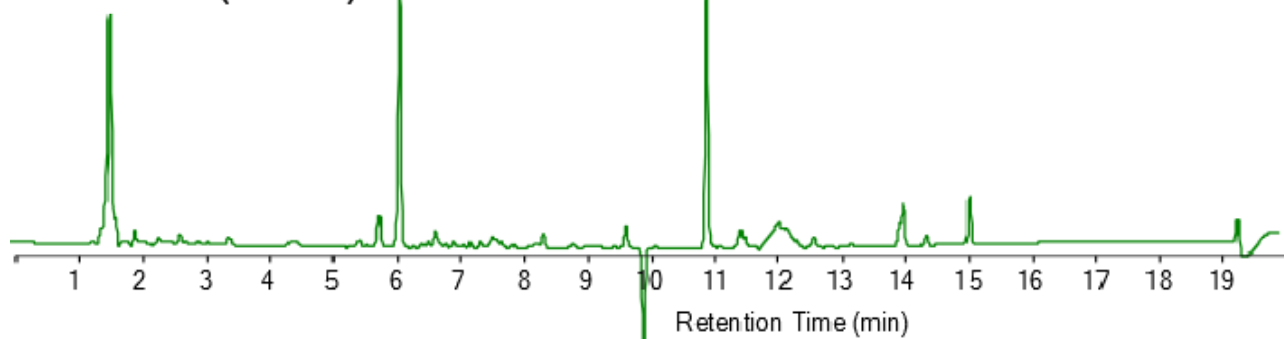
DAD1 - A:Sig=280.0,4.0 Ref=360.0,100.0 DMKP569-pos.d

AN3810 DAD



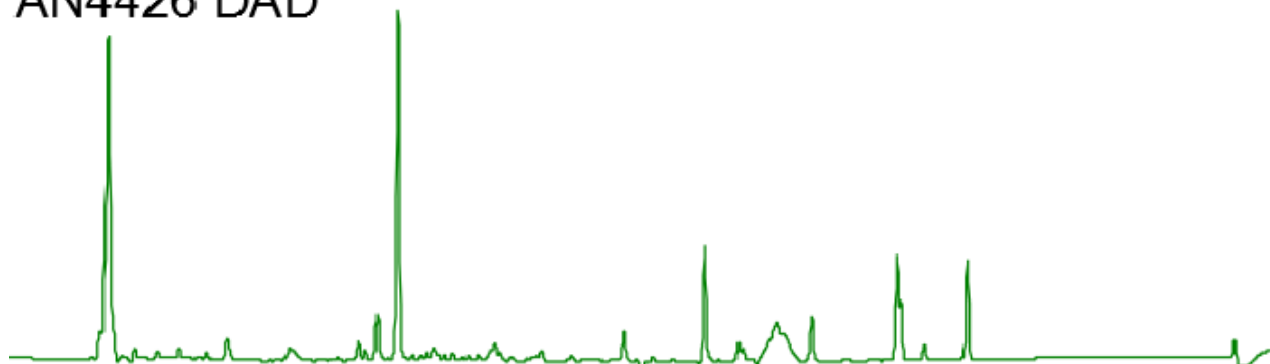
DAD1 - A:Sig=280.0,4.0 Ref=360.0,100.0 DMKP570-pos.d

AN3941 (nimT) DAD



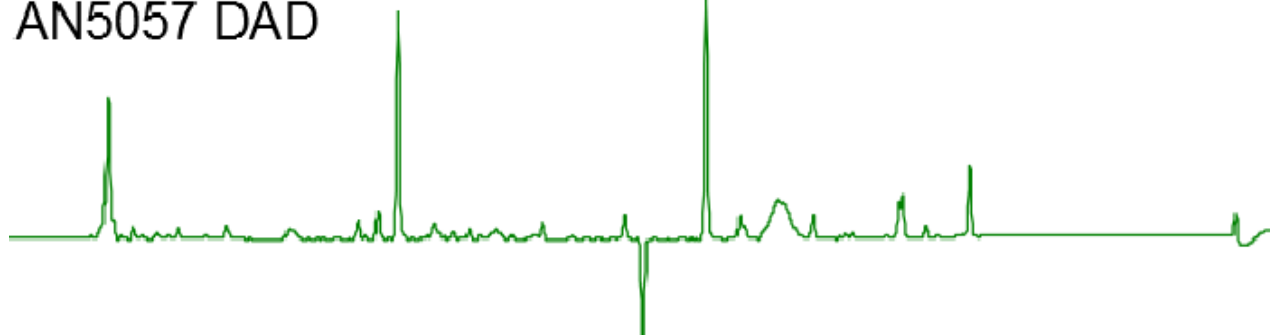
DAD1 - A:Sig=280.0,4.0 Ref=360.0,100.0 DMKP571-pos.d

AN4426 DAD



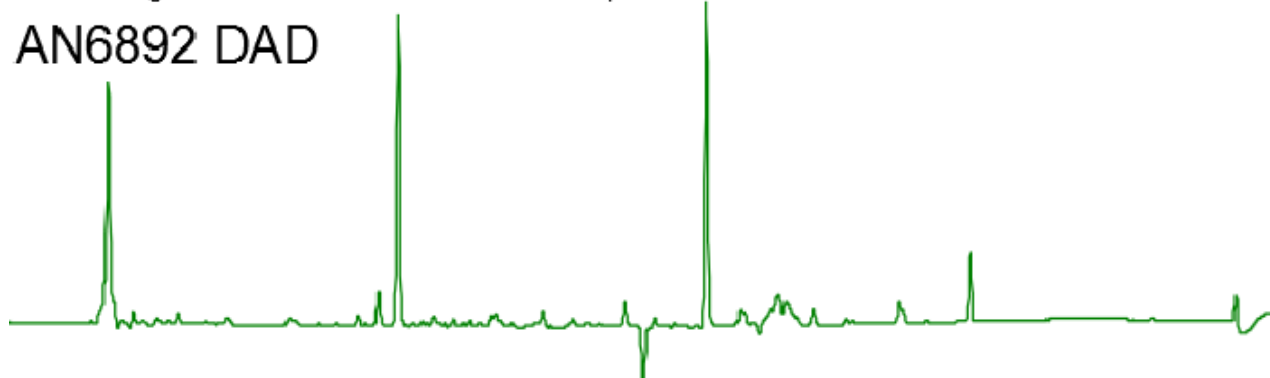
DAD1 - A:Sig=280.0,4.0 Ref=360.0,100.0 DMKP572-pos.d

AN5057 DAD



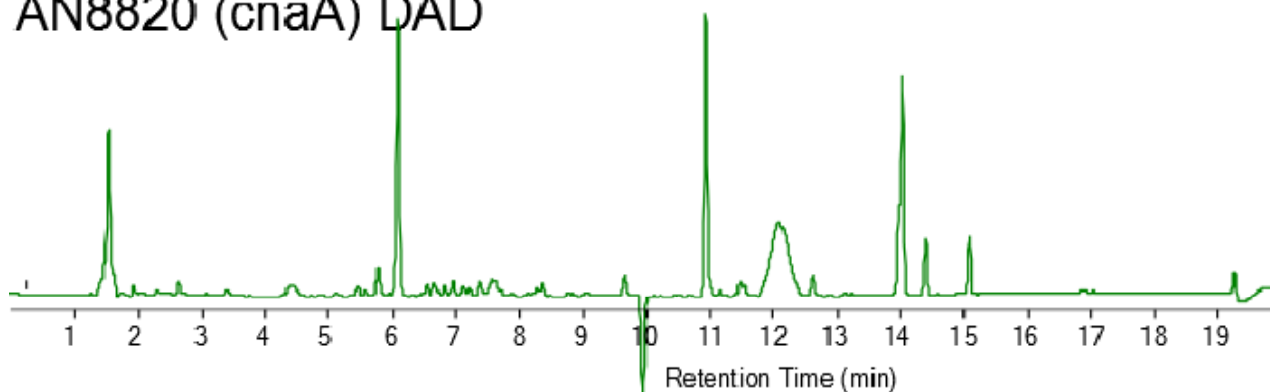
DAD1 - A:Sig=280.0,4.0 Ref=360.0,100.0 DMKP573-pos.d

AN6892 DAD



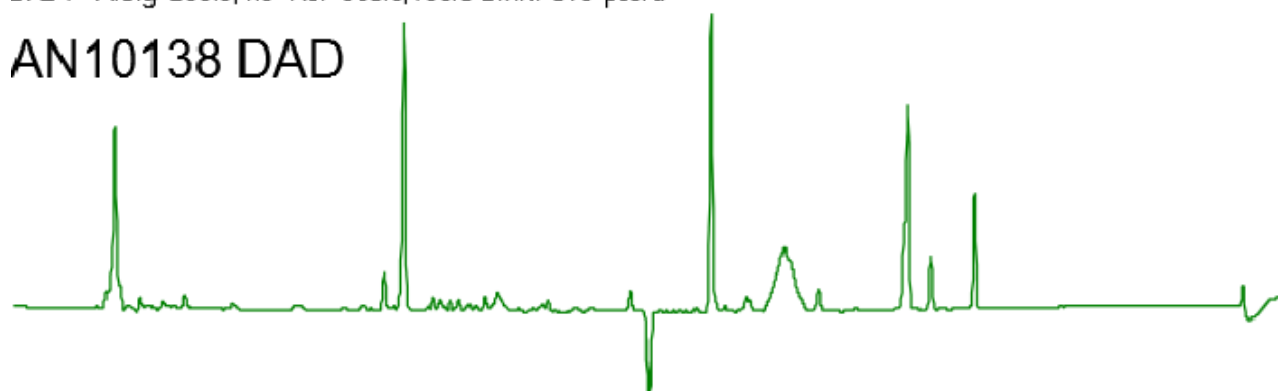
DAD1 - A:Sig=280.0,4.0 Ref=360.0,100.0 DMKP574-pos.d

AN8820 (cnaA) DAD



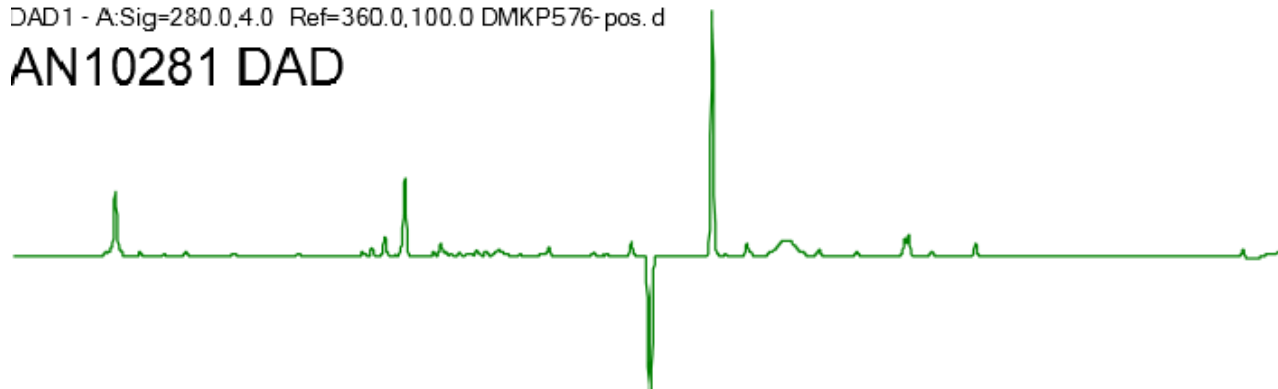
DAD1 - A:Sig=280.0,4.0 Ref=360.0,100.0 DMKP5 /5-pos. d

AN10138 DAD



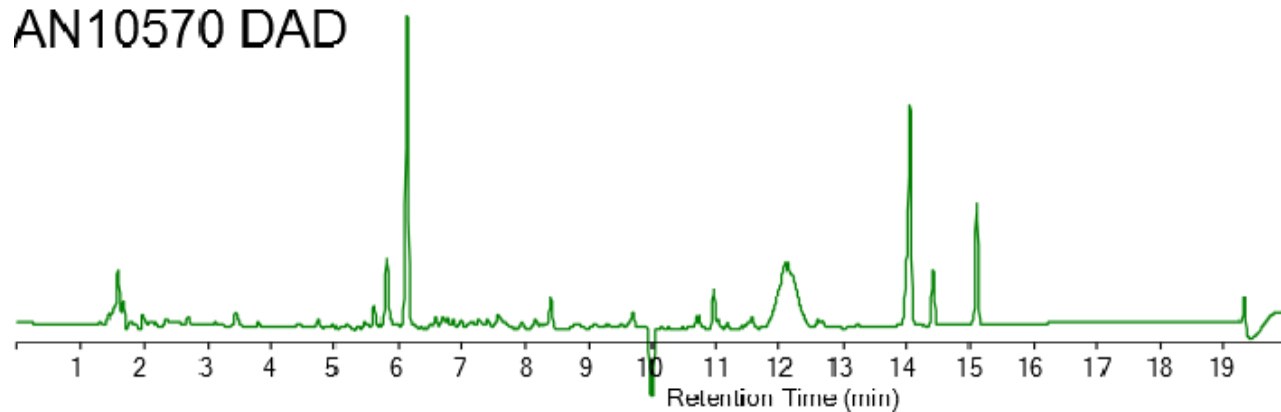
DAD1 - A:Sig=280.0,4.0 Ref=360.0,100.0 DMKP576-pos. d

AN10281 DAD



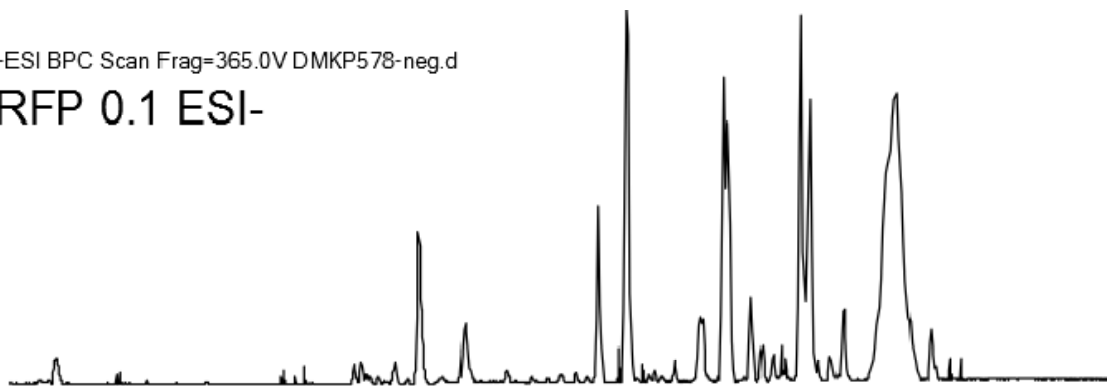
DAD1 - A:Sig=280.0,4.0 Ref=360.0,100.0 DMKP577-pos. d

AN10570 DAD



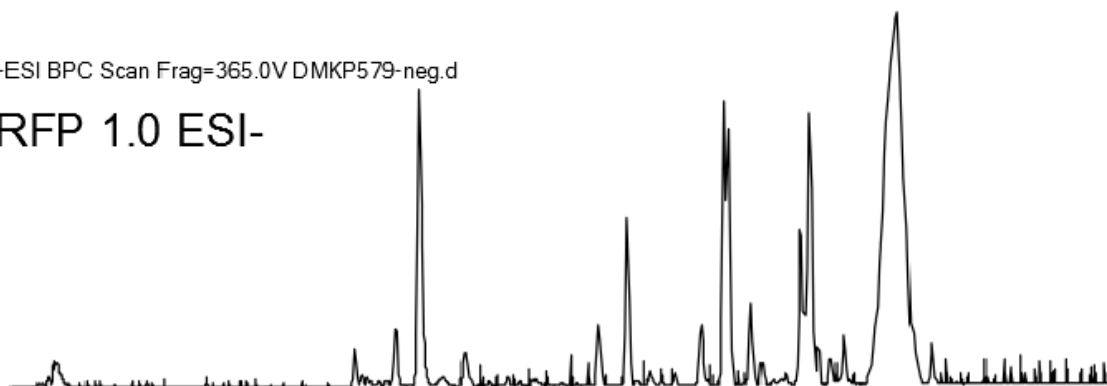
-ESI BPC Scan Frag=365.0V DMKP578-neg.d

RFP 0.1 ESI-



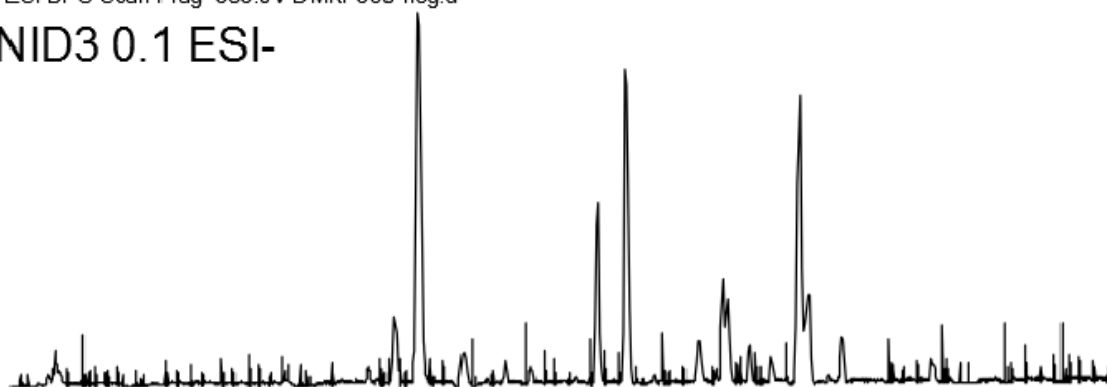
-ESI BPC Scan Frag=365.0V DMKP579-neg.d

RFP 1.0 ESI-



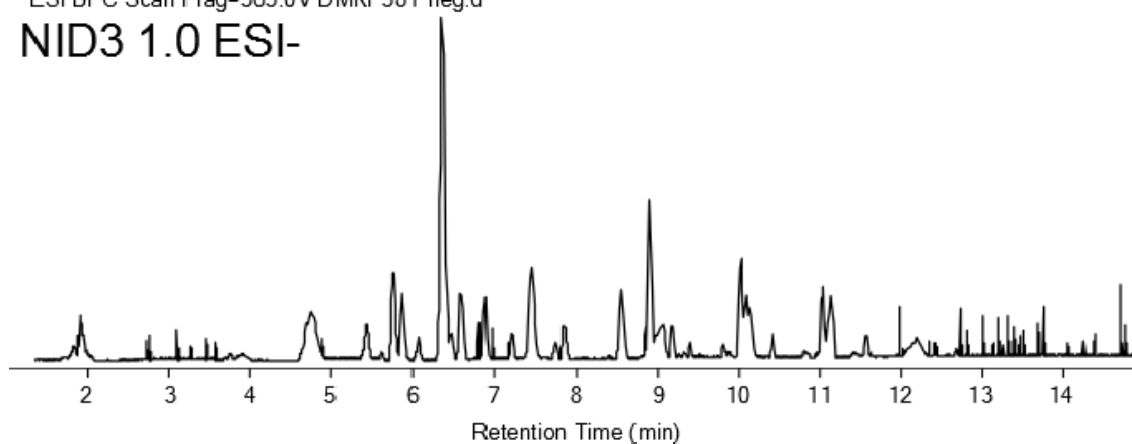
-ESI BPC Scan Frag=365.0V DMKP580-neg.d

NID3 0.1 ESI-



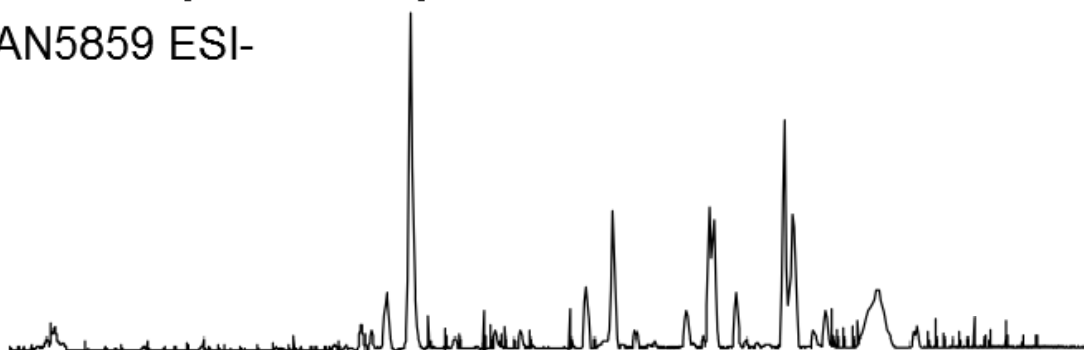
-ESI BPC Scan Frag=365.0V DMKP581-neg.d

NID3 1.0 ESI-



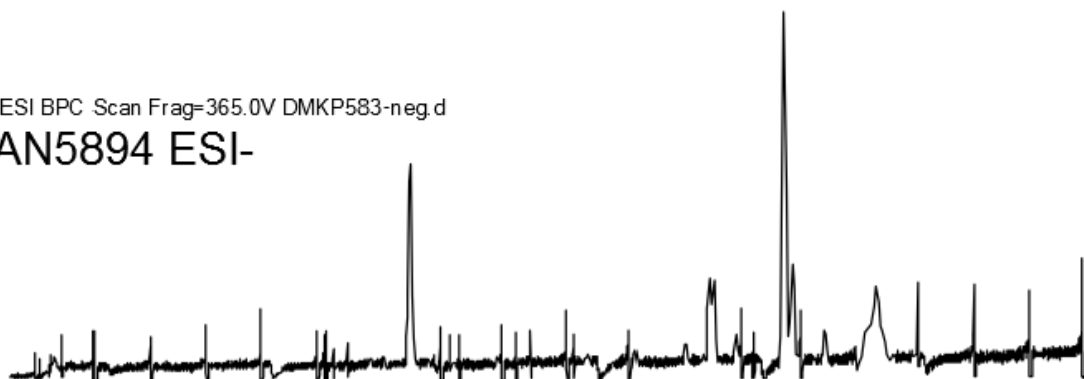
-ESI BPC Scan Frag=365.0V DMKP582-neg.d

AN5859 ESI-



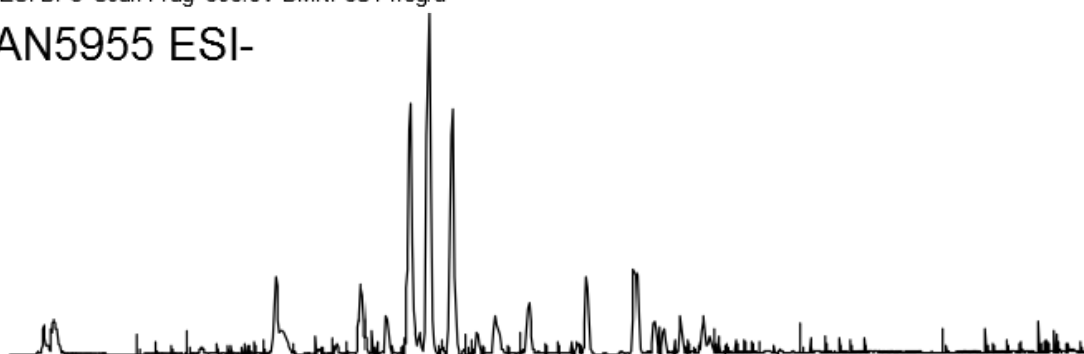
-ESI BPC Scan Frag=365.0V DMKP583-neg.d

AN5894 ESI-



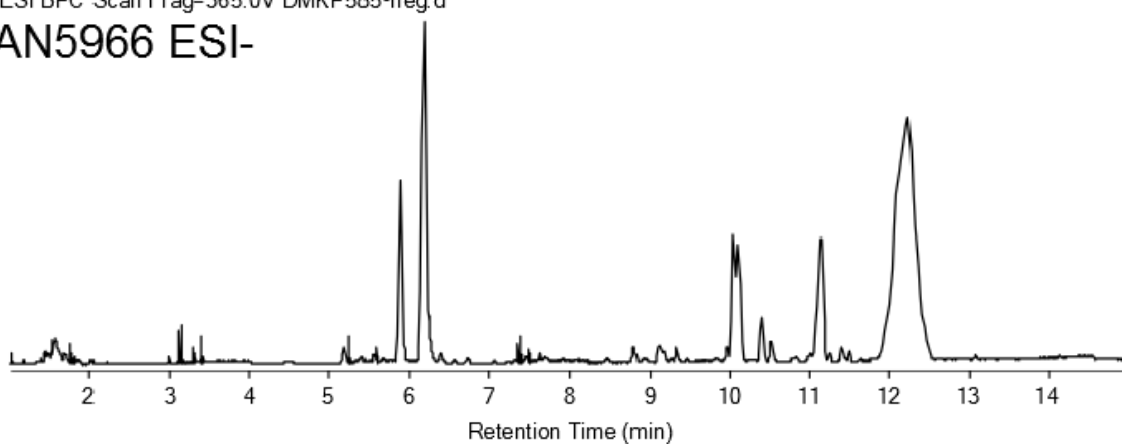
-ESI BPC Scan Frag=365.0V DMKP584-neg.d

AN5955 ESI-



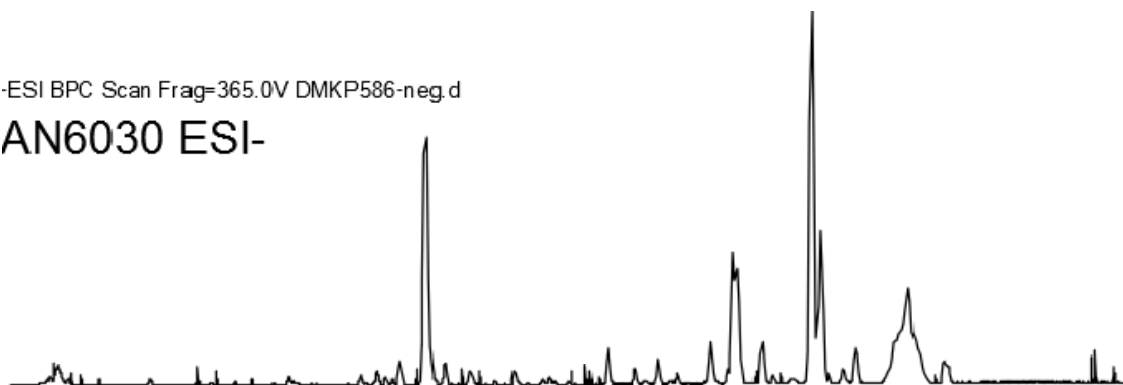
-ESI BPC Scan Frag=365.0V DMKP585-neg.d

AN5966 ESI-



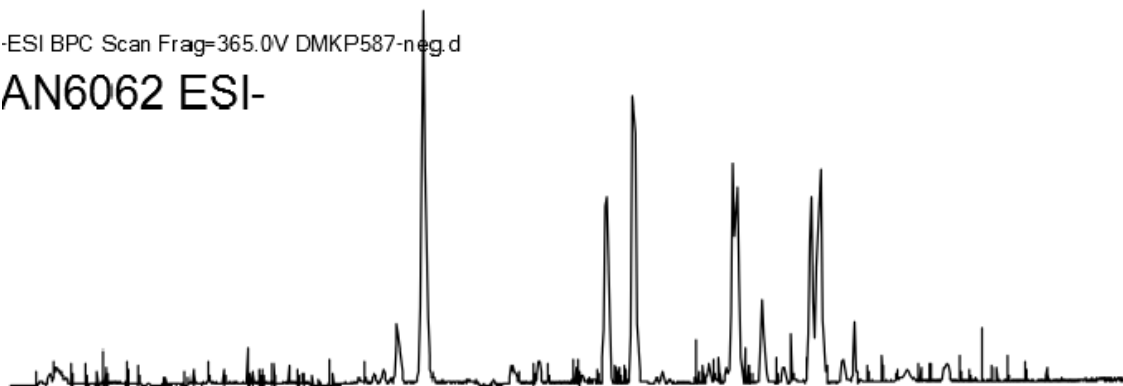
-ESI BPC Scan Frag=365.0V DMKP586-neg.d

AN6030 ESI-



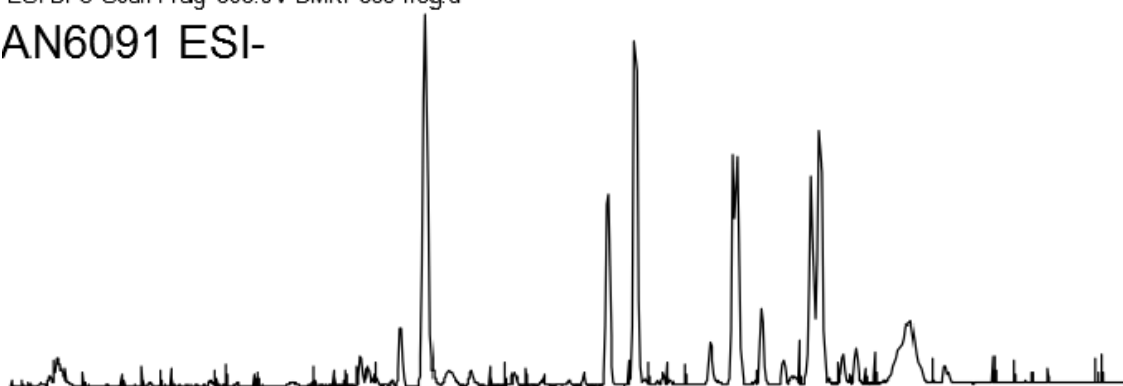
-ESI BPC Scan Frag=365.0V DMKP587-neg.d

AN6062 ESI-



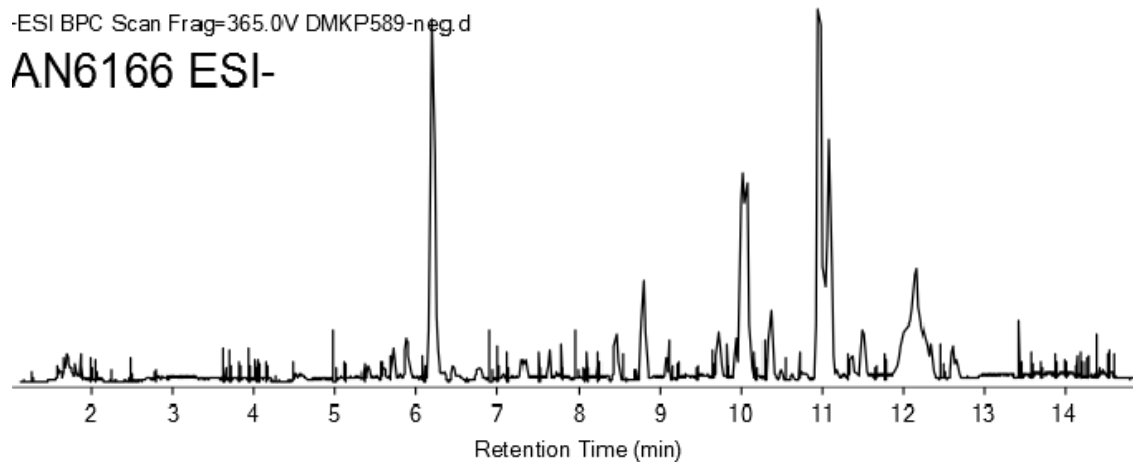
-ESI BPC Scan Frag=365.0V DMKP588-neg.d

AN6091 ESI-



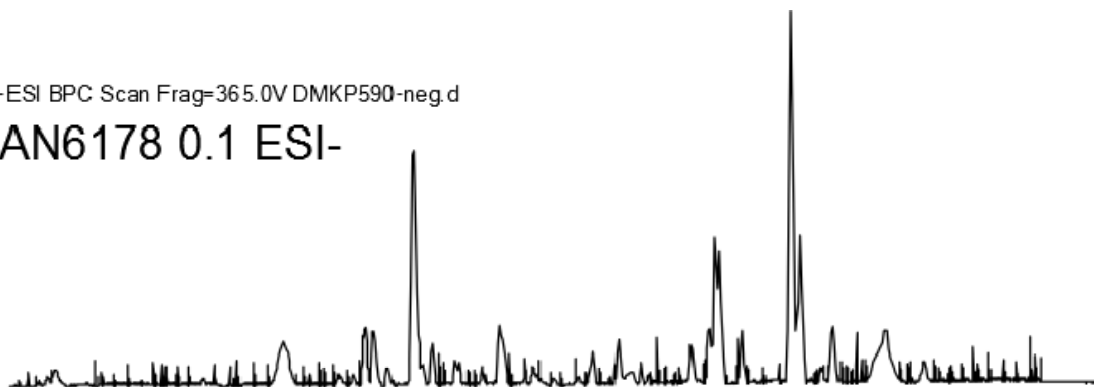
-ESI BPC Scan Frag=365.0V DMKP589-neg.d

AN6166 ESI-



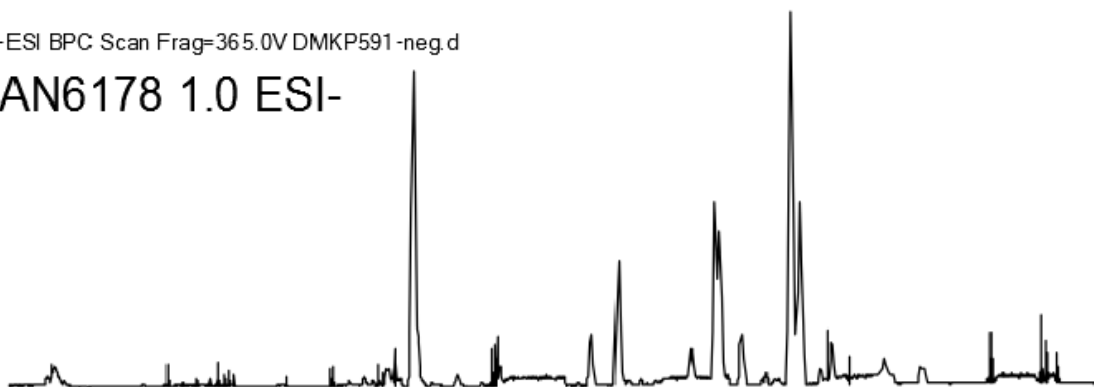
-ESI BPC Scan Frag=365.0V DMKP590-neg.d

AN6178 0.1 ESI-



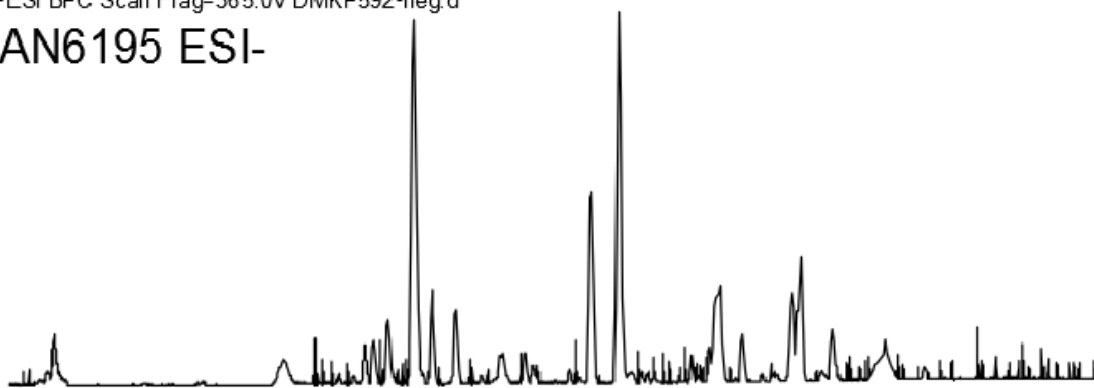
-ESI BPC Scan Frag=365.0V DMKP591-neg.d

AN6178 1.0 ESI-



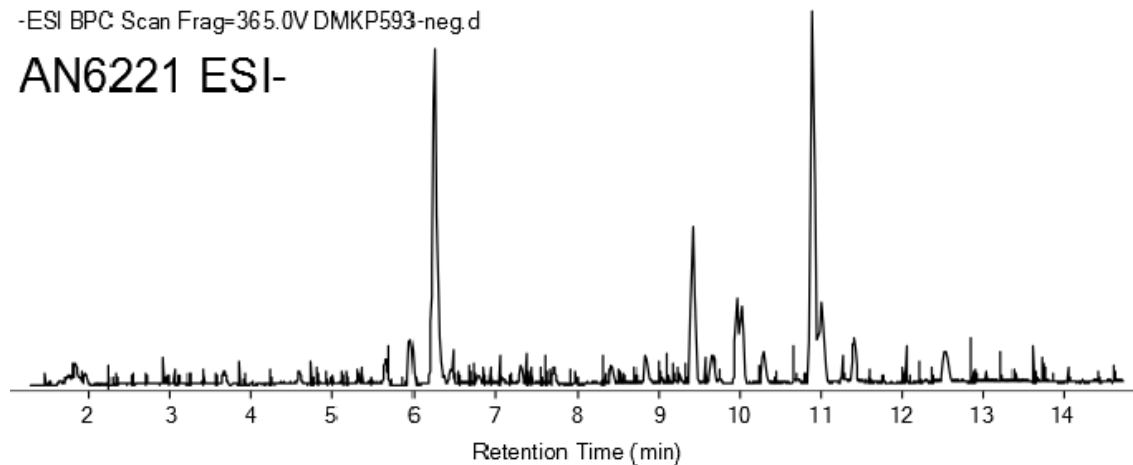
-ESI BPC Scan Frag=365.0V DMKP592-neg.d

AN6195 ESI-



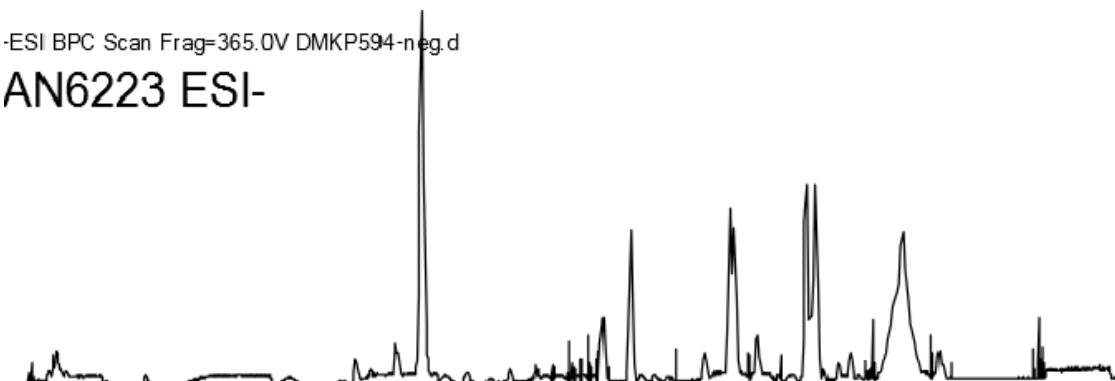
-ESI BPC Scan Frag=365.0V DMKP593-neg.d

AN6221 ESI-



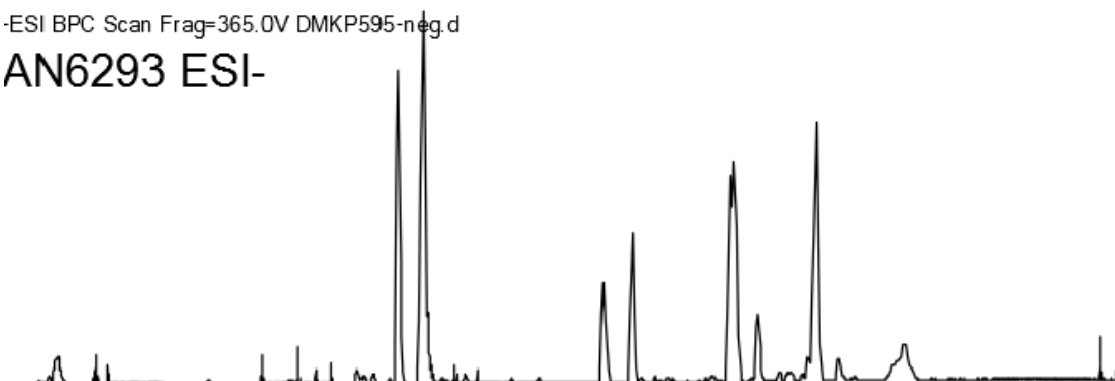
-ESI BPC Scan Frag=365.0V DMKP594-neg.d

AN6223 ESI-



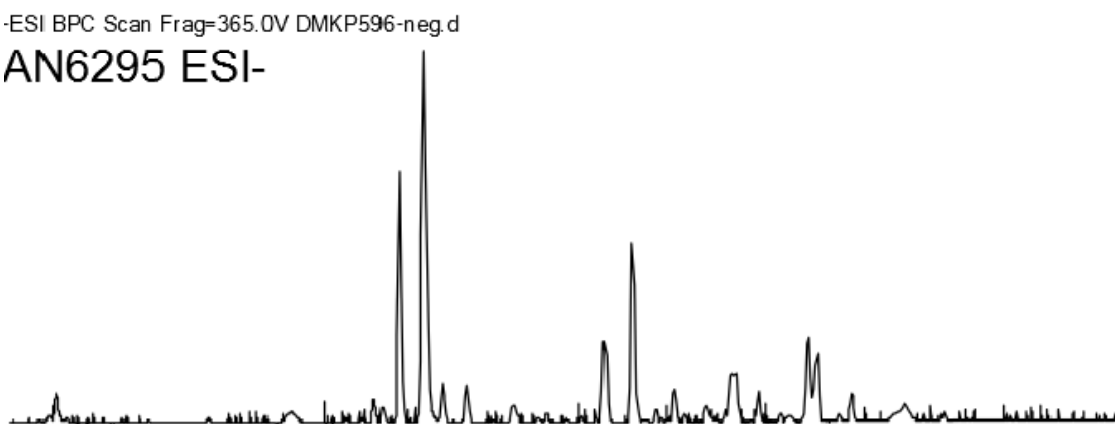
-ESI BPC Scan Frag=365.0V DMKP595-neg.d

AN6293 ESI-



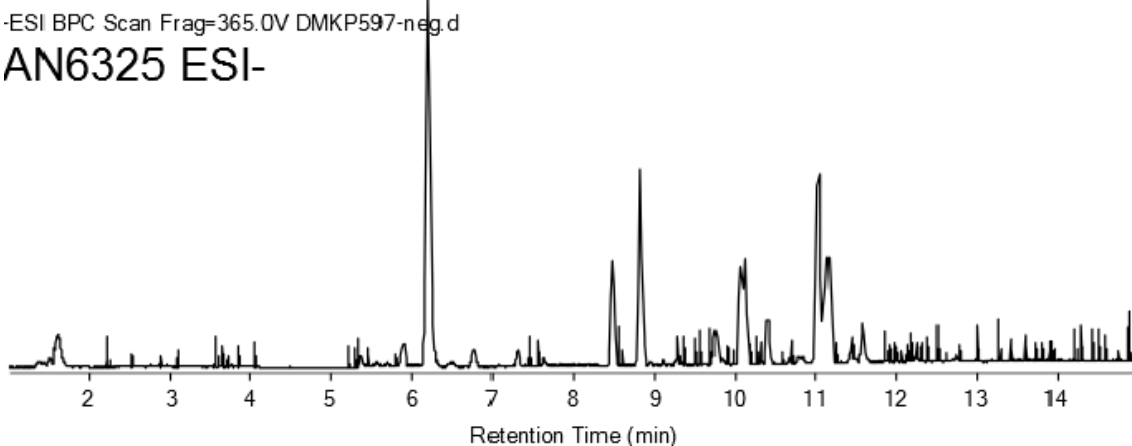
-ESI BPC Scan Frag=365.0V DMKP596-neg.d

AN6295 ESI-



-ESI BPC Scan Frag=365.0V DMKP597-neg.d

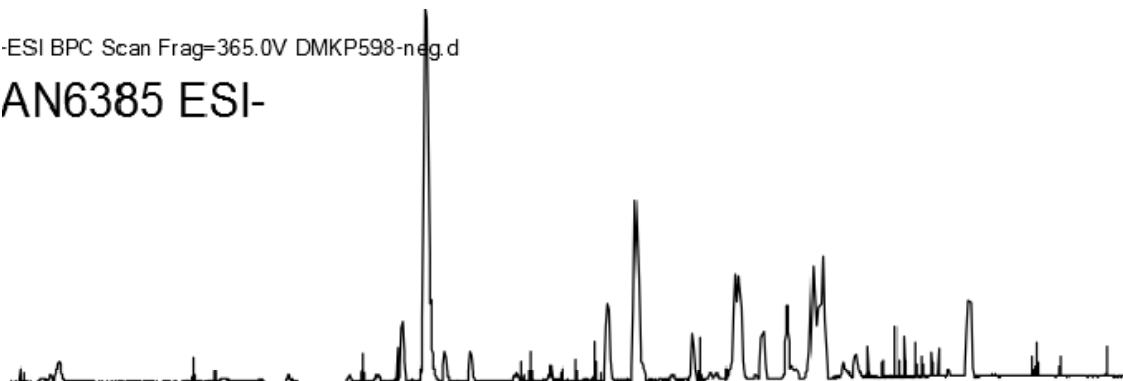
AN6325 ESI-



Retention Time (min)

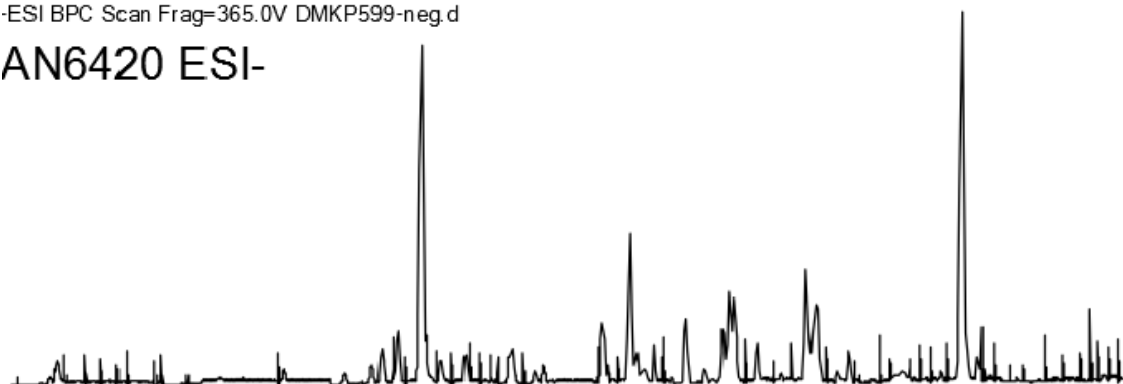
-ESI BPC Scan Frag=365.0V DMKP598-neg.d

AN6385 ESI-



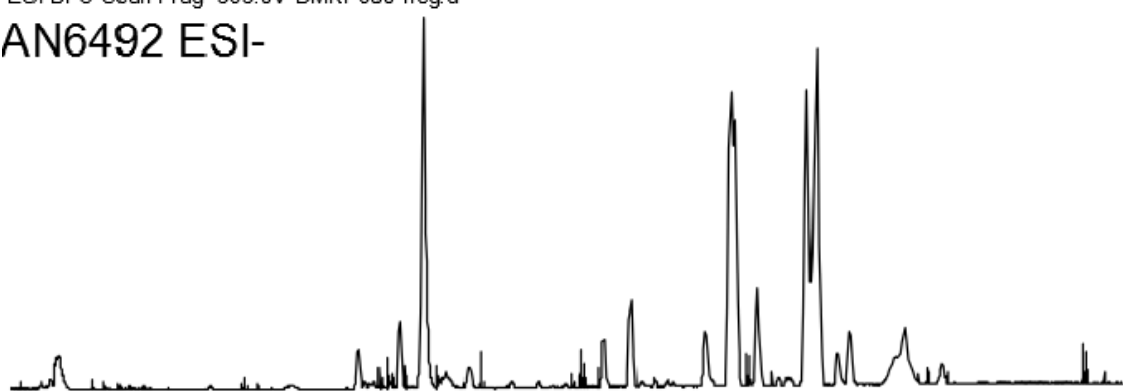
-ESI BPC Scan Frag=365.0V DMKP599-neg.d

AN6420 ESI-



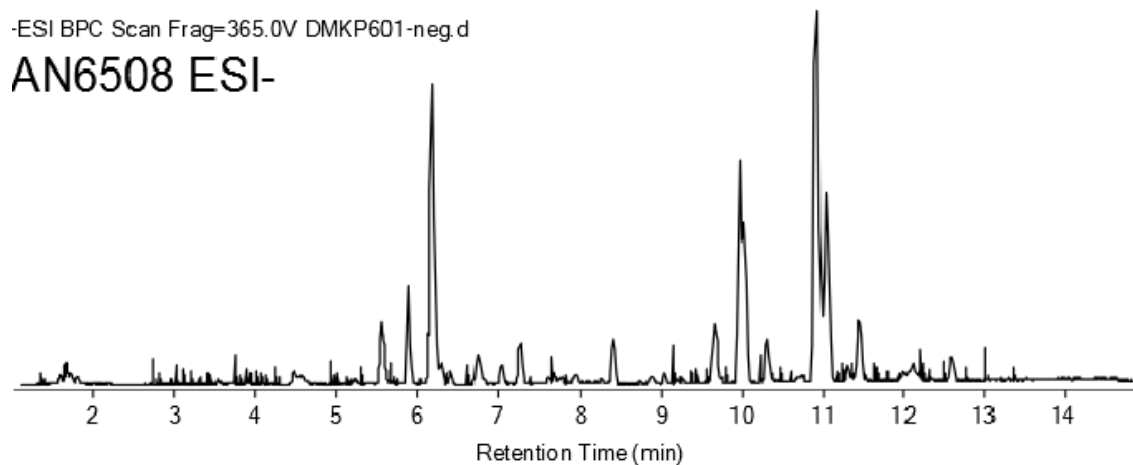
-ESI BPC Scan Frag=365.0V DMKP600-neg.d

AN6492 ESI-



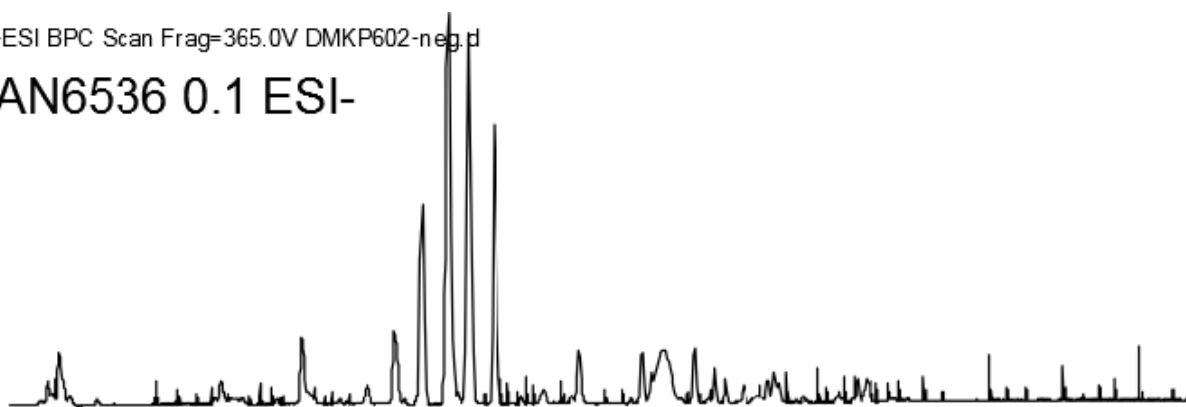
-ESI BPC Scan Frag=365.0V DMKP601-neg.d

AN6508 ESI-



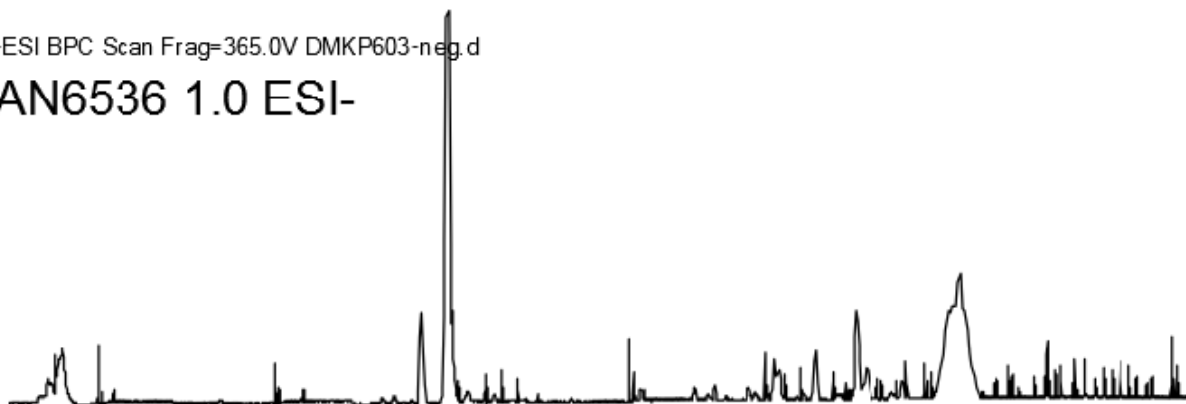
-ESI BPC Scan Frag=365.0V DMKP602-neg.d

AN6536 0.1 ESI-



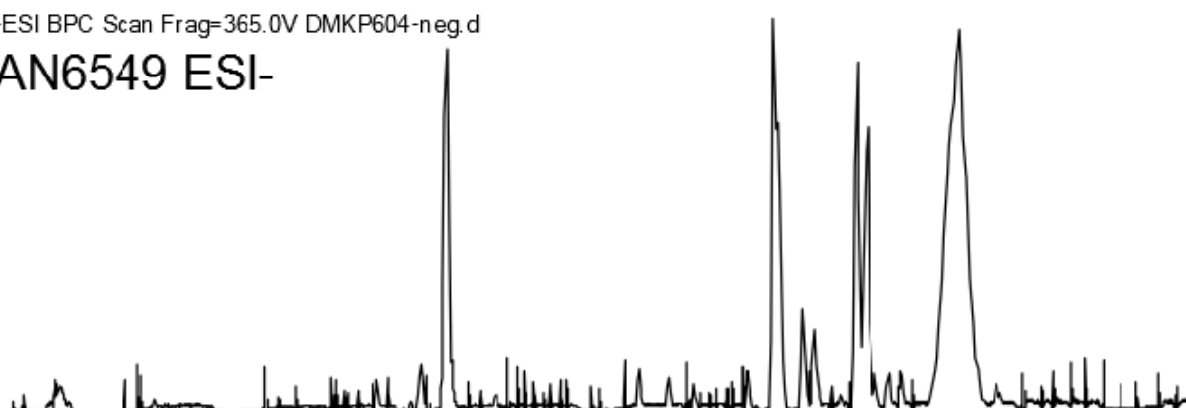
-ESI BPC Scan Frag=365.0V DMKP603-neg.d

AN6536 1.0 ESI-



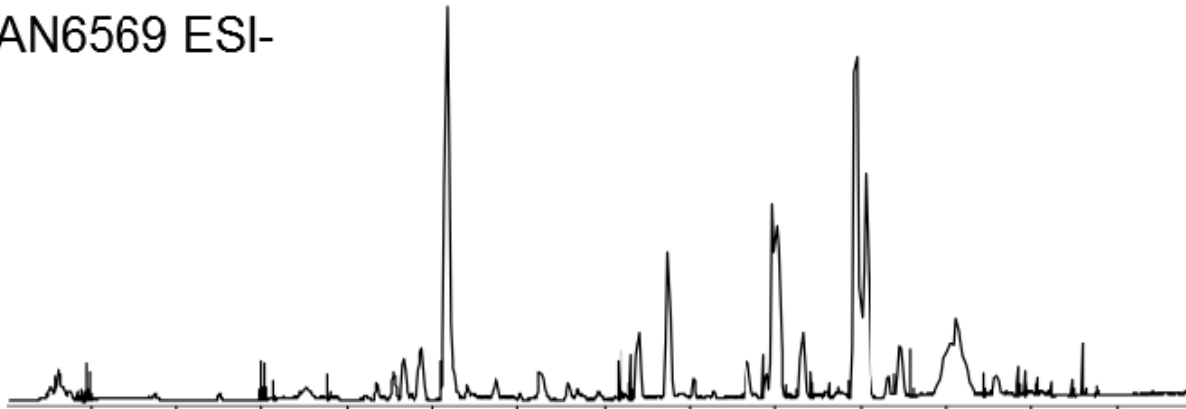
-ESI BPC Scan Frag=365.0V DMKP604-neg.d

AN6549 ESI-



-ESI BPC Scan Frag=365.0V DMKP605-neg.d

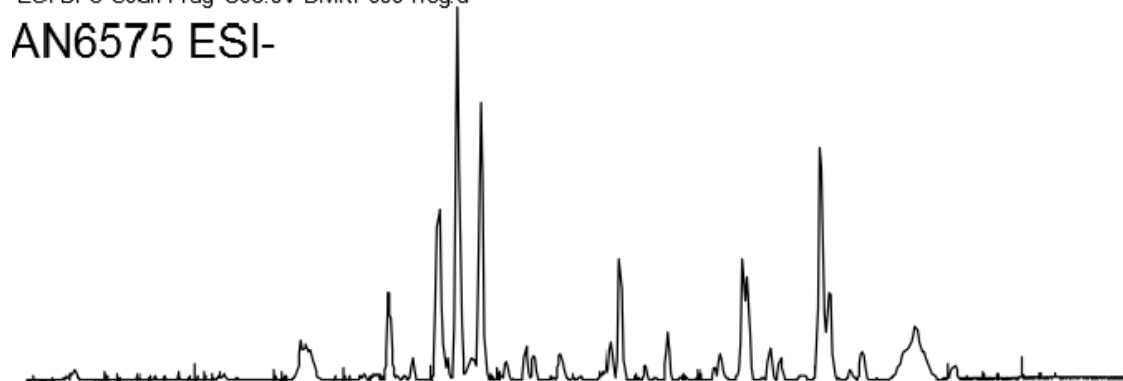
AN6569 ESI-



Retention Time (min)

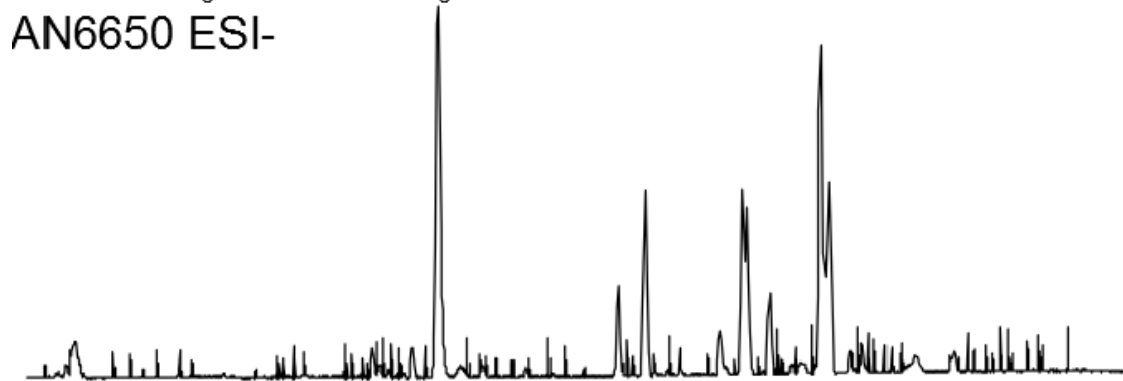
-ESI BPC Scan Frag=365.0V DMKP606-neg.d

AN6575 ESI-



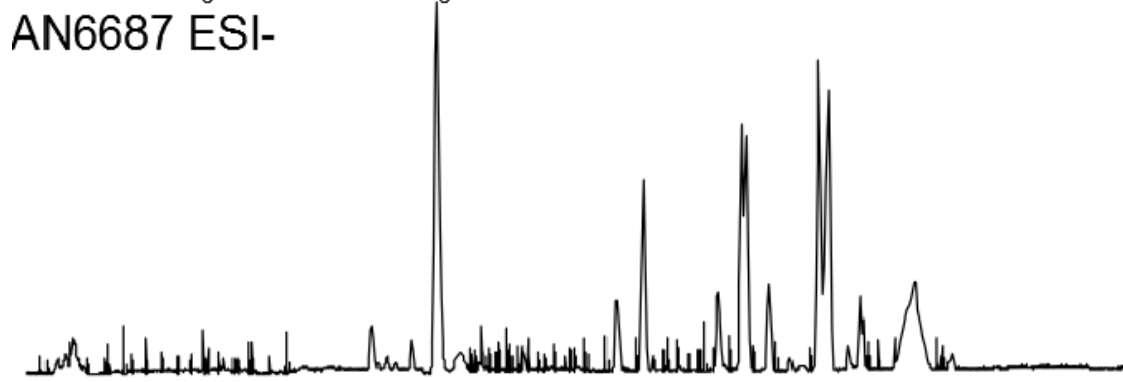
-ESI BPC Scan Frag=365.0V DMKP607-neg.d

AN6650 ESI-



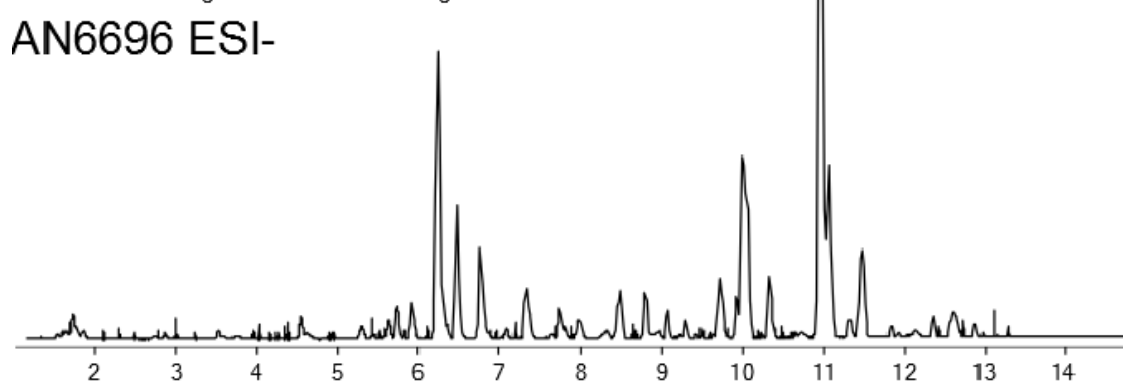
-ESI BPC Scan Frag=365.0V DMKP608-neg.d

AN6687 ESI-



-ESI BPC Scan Frag=365.0V DMKP609-neg.d

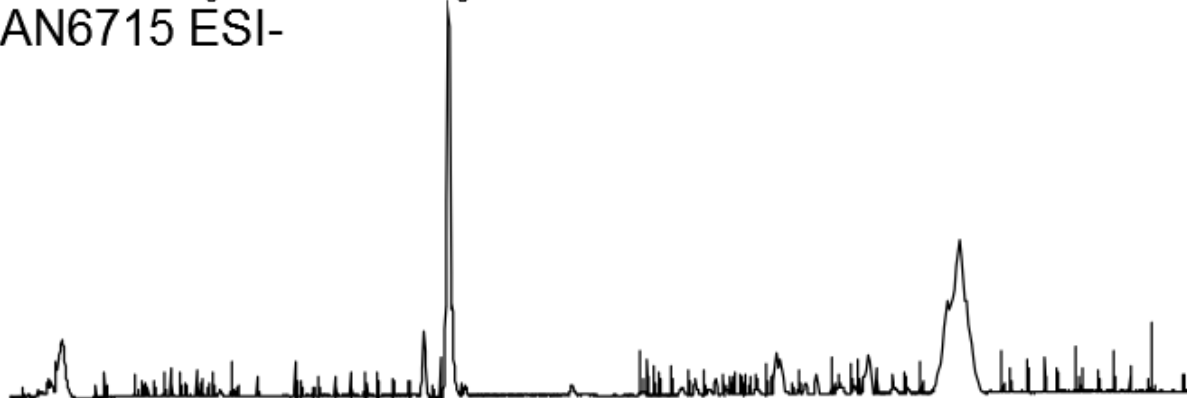
AN6696 ESI-



Retention Time (min)

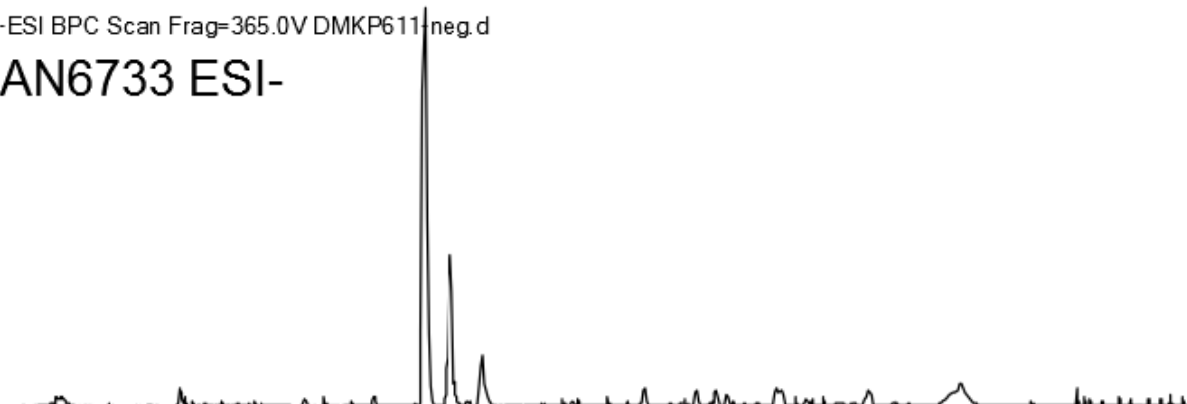
-ESI BPC Scan Frag=365.0V DMKP610-neg.d

AN6715 ESI-



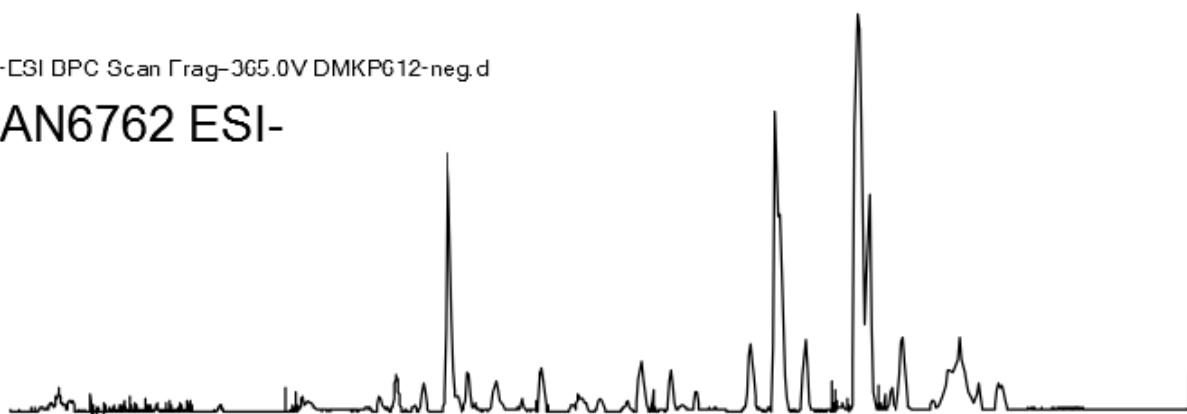
-ESI BPC Scan Frag=365.0V DMKP611-neg.d

AN6733 ESI-



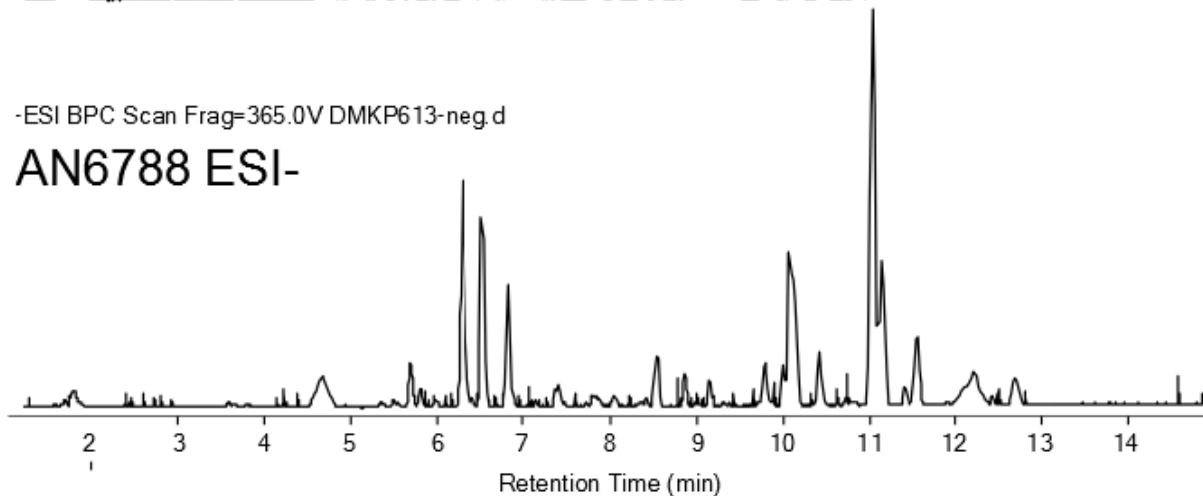
-ESI BPC Scan Frag=365.0V DMKP612-neg.d

AN6762 ESI-



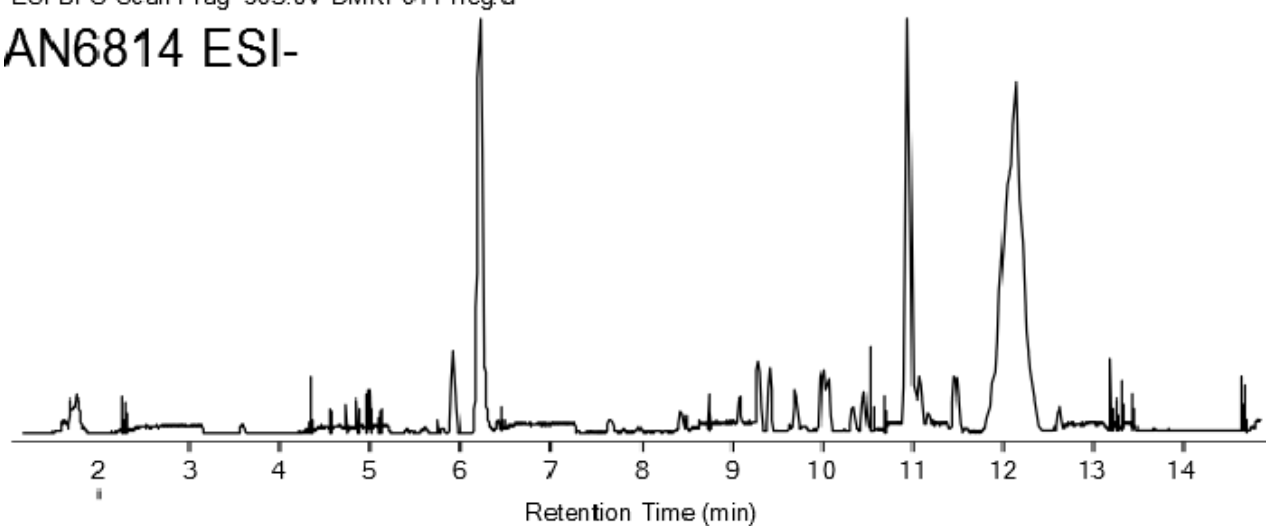
-ESI BPC Scan Frag=365.0V DMKP613-neg.d

AN6788 ESI-



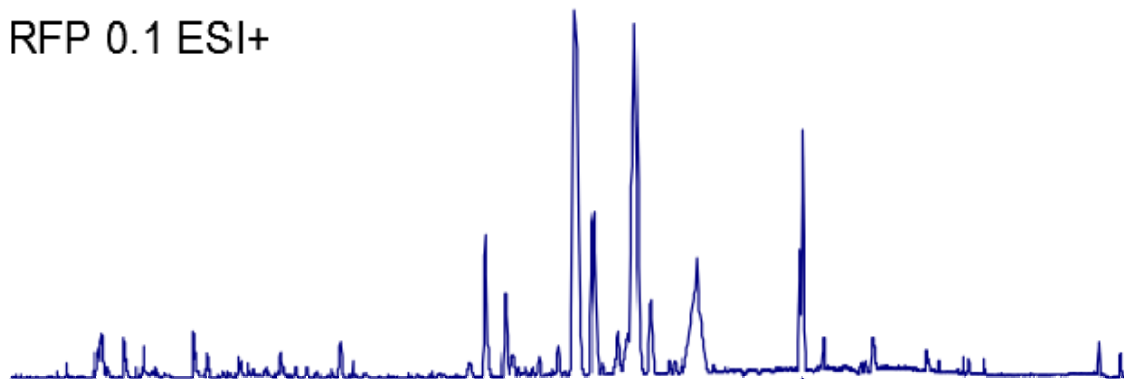
-ESI BPC Scan Frag=365.0V DMKP614-neg.d

AN6814 ESI-



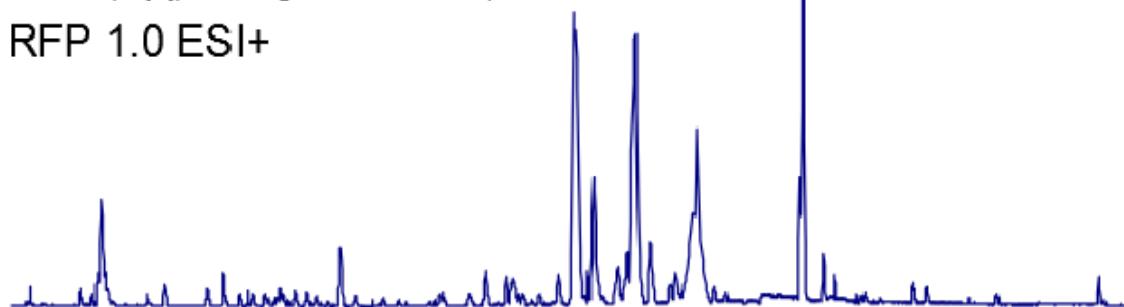
+ESI BPC(all [-1]) Scan Frag=365.0V DMKP578-pos.d

RFP 0.1 ESI+



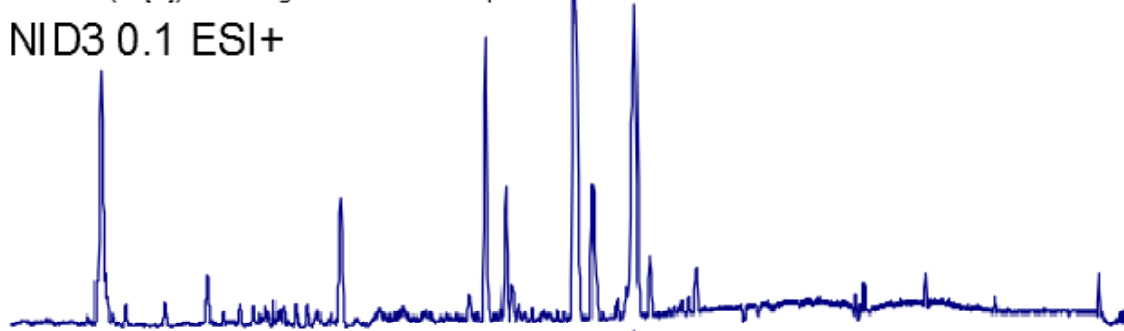
+ESI BPC(all [-1]) Scan Frag=365.0V DMKP579-pos.d

RFP 1.0 ESI+



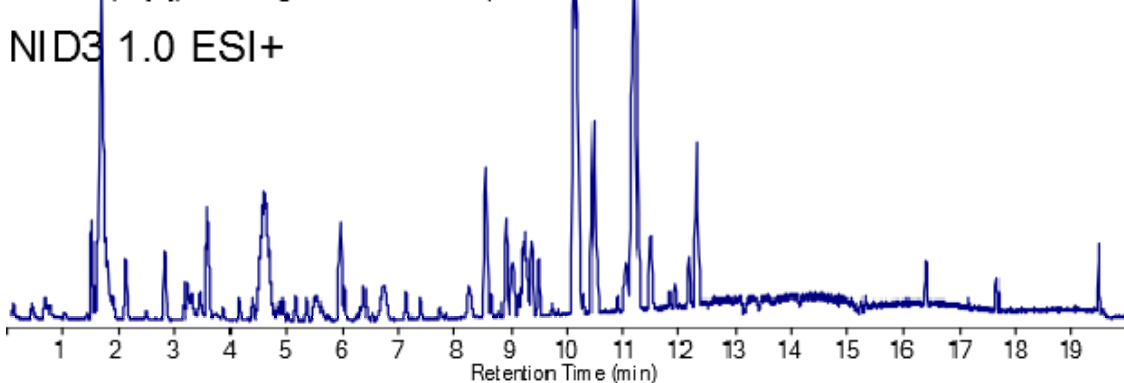
+ESI BPC(all [-1]) Scan Frag=365.0V DMKP580-pos.d

NID3 0.1 ESI+



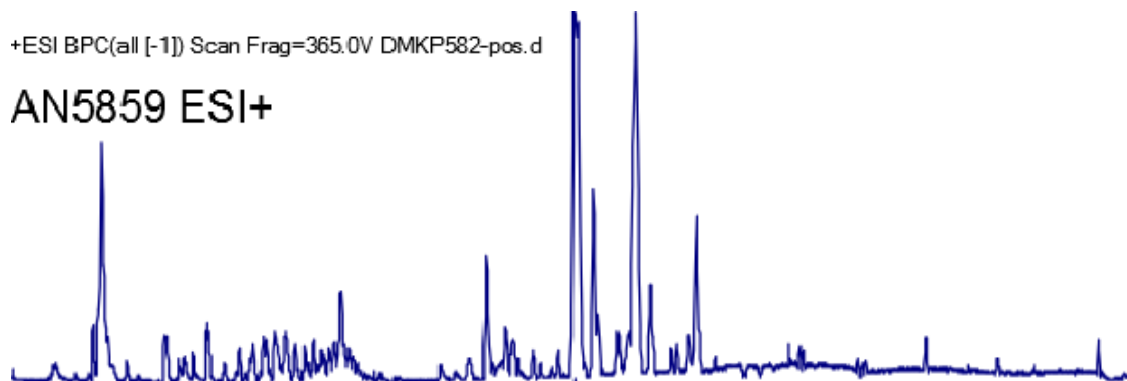
+ESI BPC(all [-1]) Scan Frag=365.0V DMKP581-pos.d

NID3 1.0 ESI+



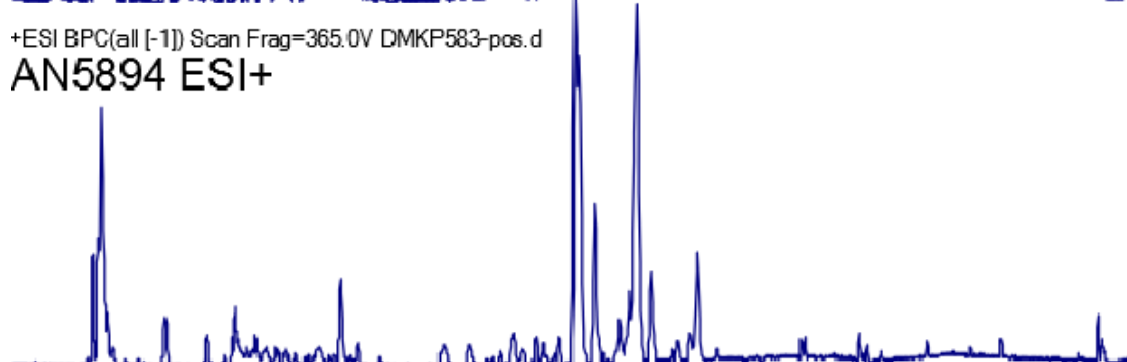
+ESI BPC(all [-1]) Scan Frag=365.0V DMKP582-pos.d

AN5859 ESI+



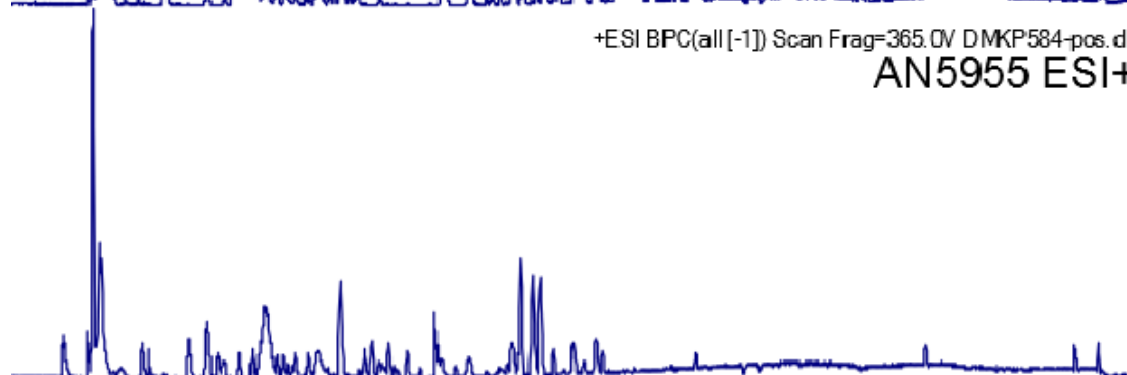
+ESI BPC(all [-1]) Scan Frag=365.0V DMKP583-pos.d

AN5894 ESI+



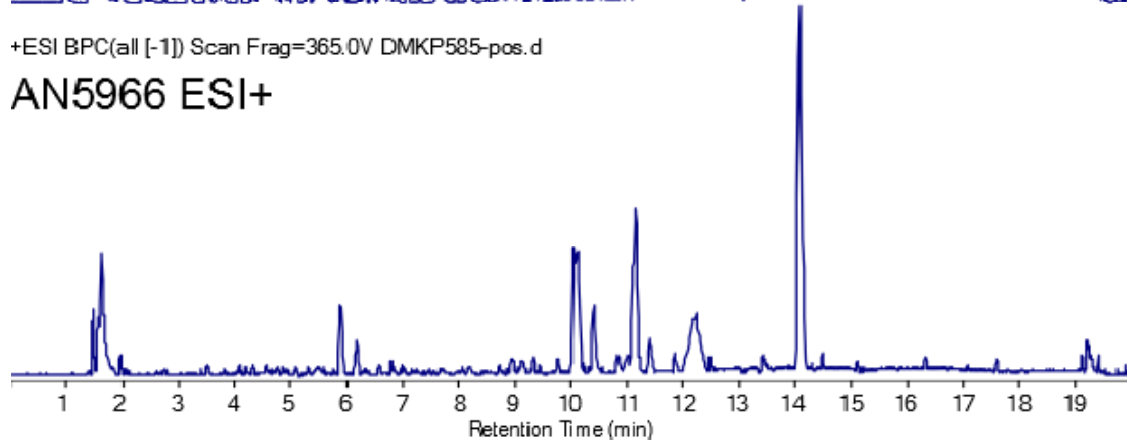
+ESI BPC(all [-1]) Scan Frag=365.0V DMKP584-pos.d

AN5955 ESI+



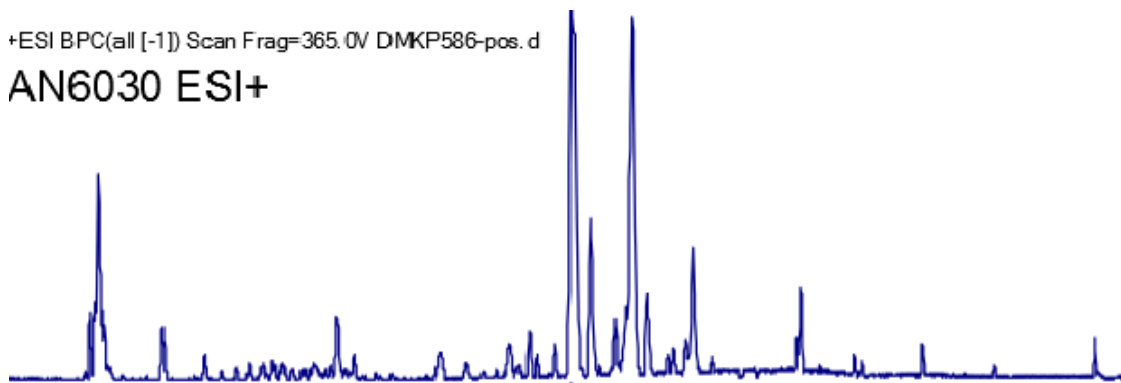
+ESI BPC(all [-1]) Scan Frag=365.0V DMKP585-pos.d

AN5966 ESI+



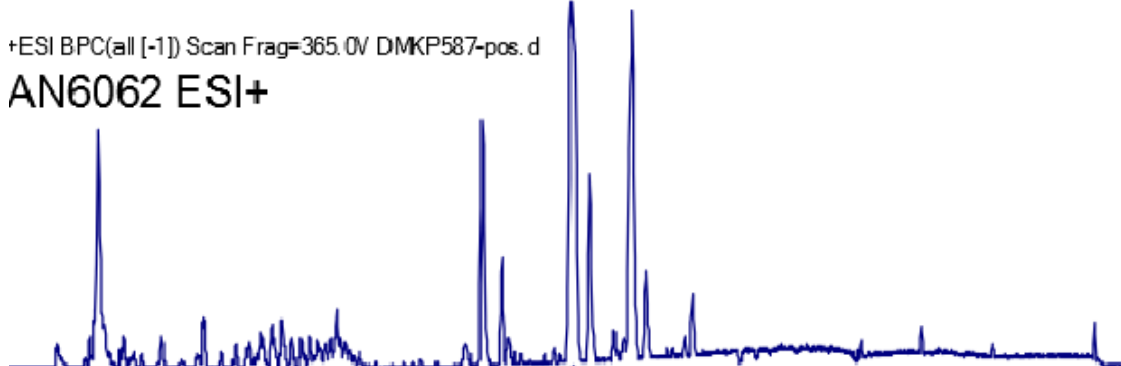
+ESI BPC(all [-1]) Scan Frag=365.0V DMKP586-pos. d

AN6030 ESI+



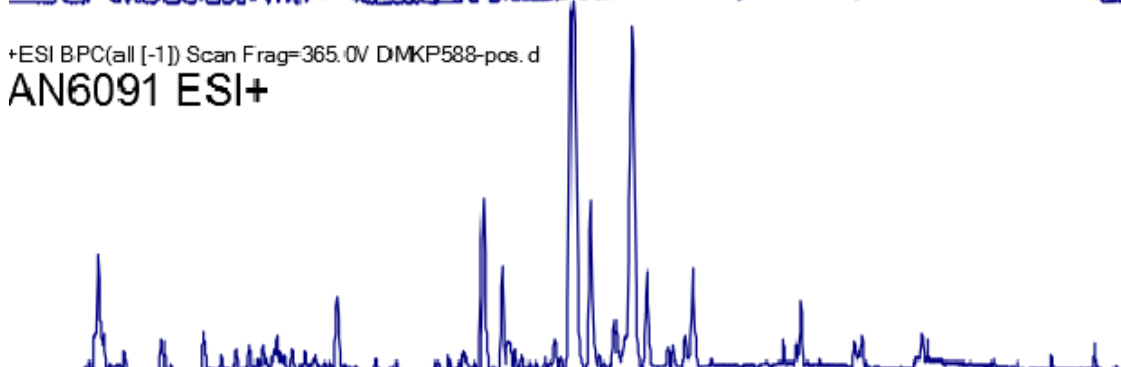
+ESI BPC(all [-1]) Scan Frag=365.0V DMKP587-pos. d

AN6062 ESI+



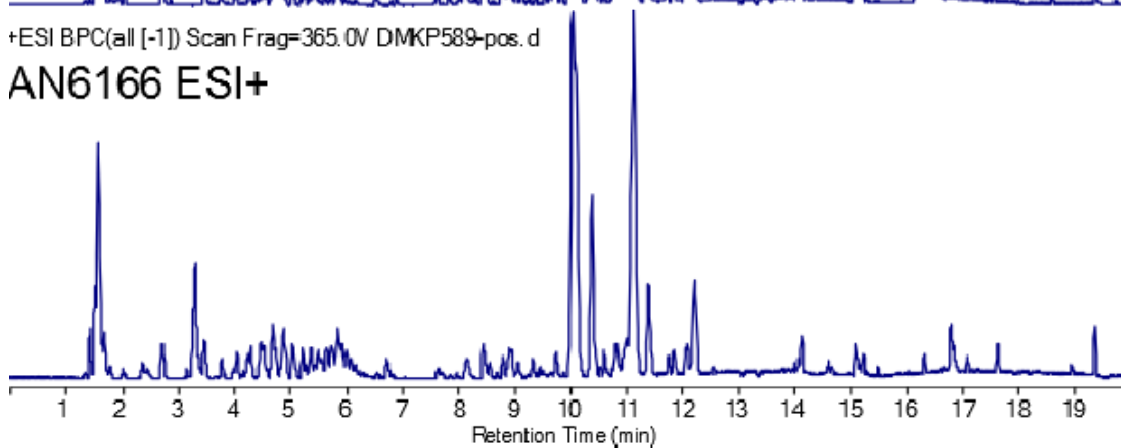
+ESI BPC(all [-1]) Scan Frag=365.0V DMKP588-pos. d

AN6091 ESI+



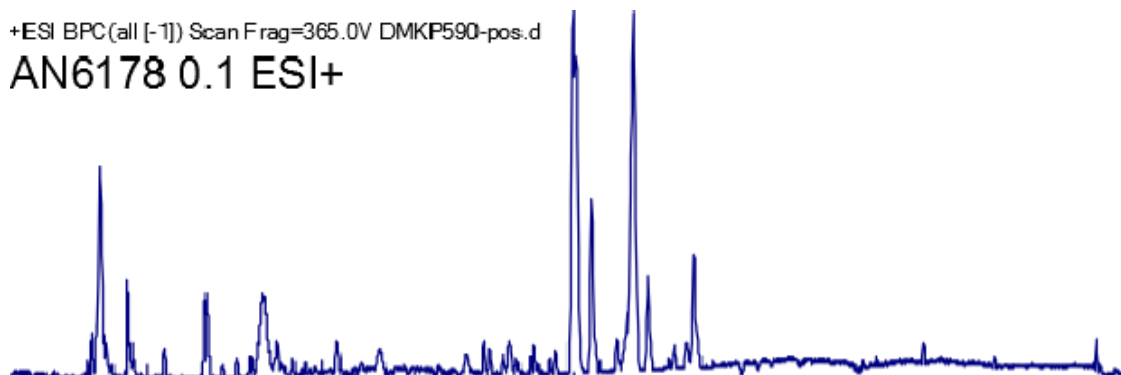
+ESI BPC(all [-1]) Scan Frag=365.0V DMKP589-pos. d

AN6166 ESI+



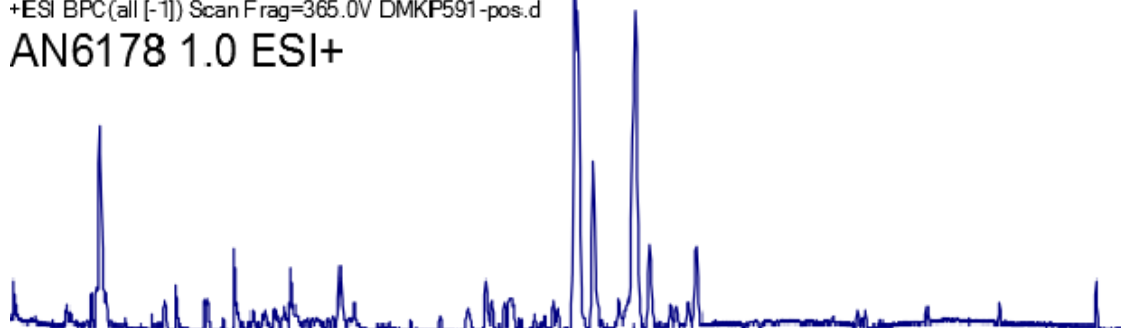
+ESI BPC(all [-1]) Scan Frag=365.0V DMKP590-pos.d

AN6178 0.1 ESI+



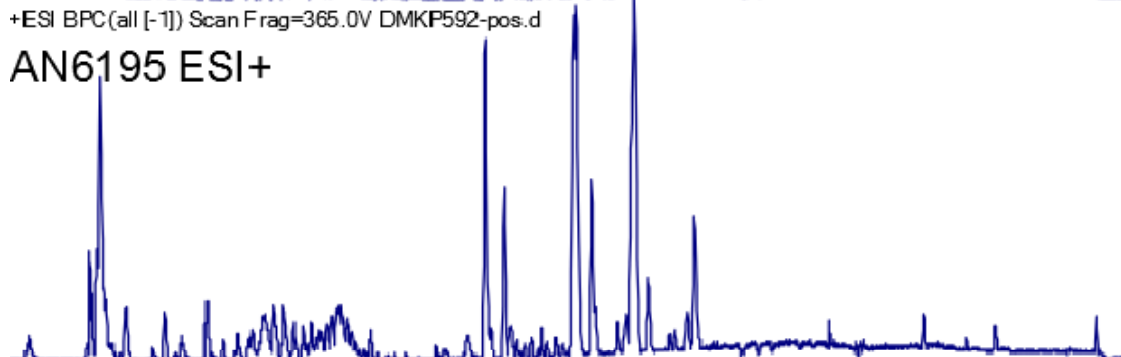
+ESI BPC(all [-1]) Scan Frag=365.0V DMKP591-pos.d

AN6178 1.0 ESI+



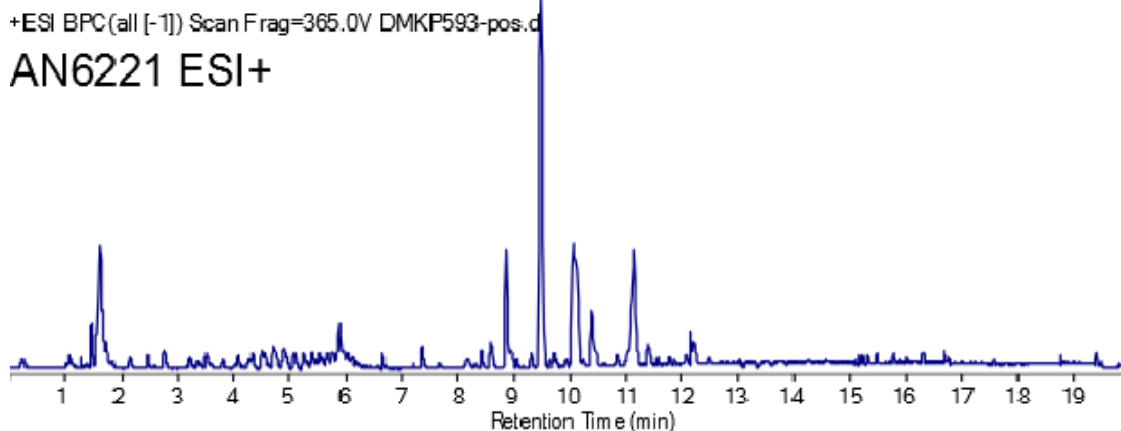
+ESI BPC(all [-1]) Scan Frag=365.0V DMKP592-pos.d

AN6195 ESI+



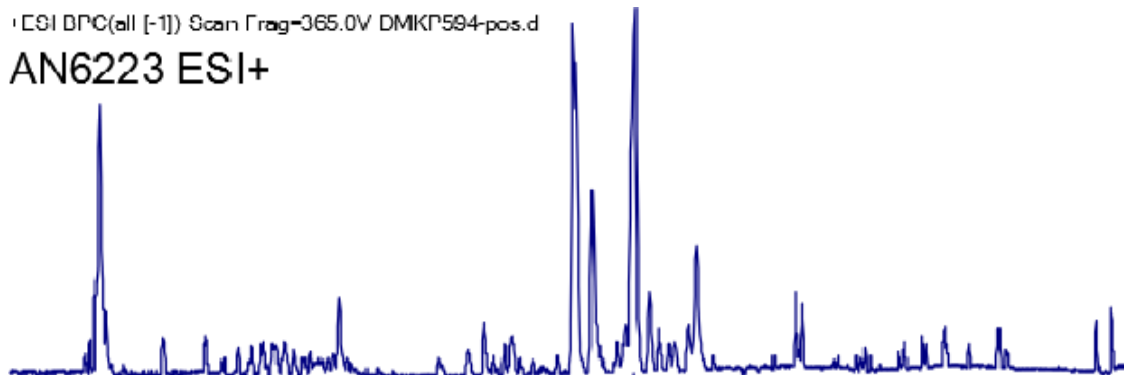
+ESI BPC(all [-1]) Scan Frag=365.0V DMKP593-pos.d

AN6221 ESI+



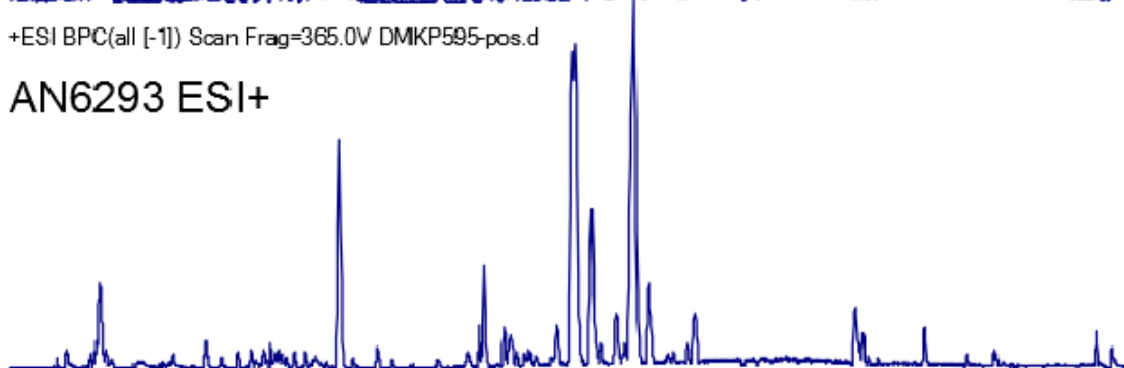
+ESI BPC(all [-1]) Scan Frag=365.0V DMKP594-pos.d

AN6223 ESI+



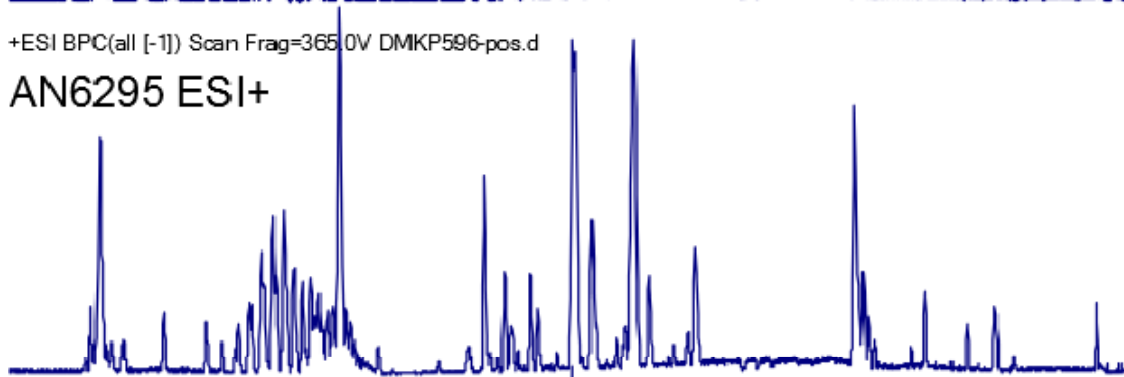
+ESI BPC(all [-1]) Scan Frag=365.0V DMKP595-pos.d

AN6293 ESI+



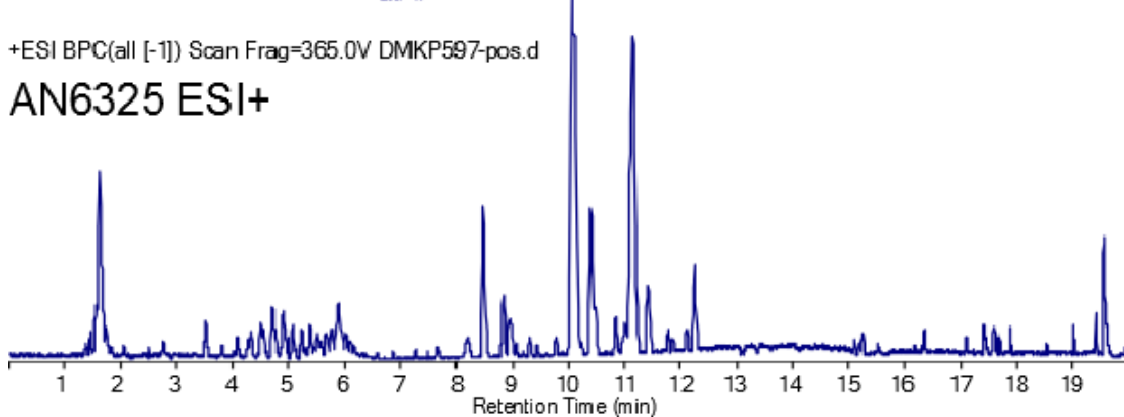
+ESI BPC(all [-1]) Scan Frag=365.0V DMKP596-pos.d

AN6295 ESI+



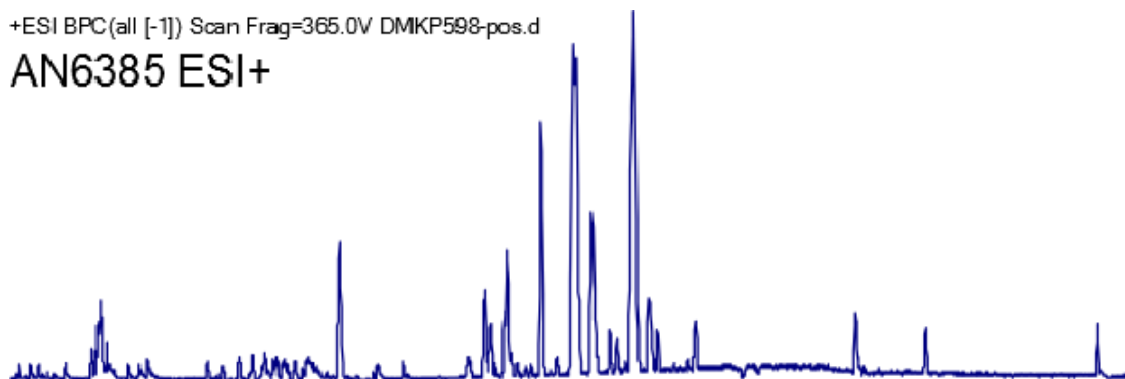
+ESI BPC(all [-1]) Scan Frag=365.0V DMKP597-pos.d

AN6325 ESI+



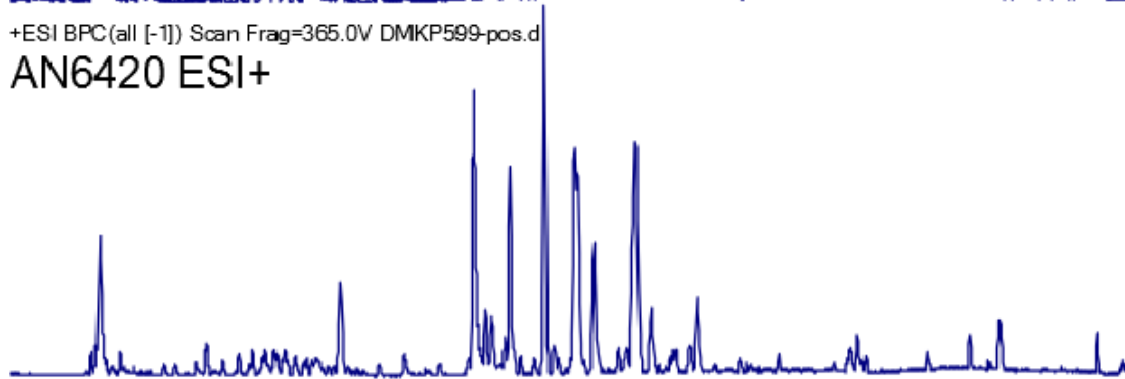
+ESI BPC(all [-1]) Scan Frag=365.0V DMKP598-pos.d

AN6385 ESI+



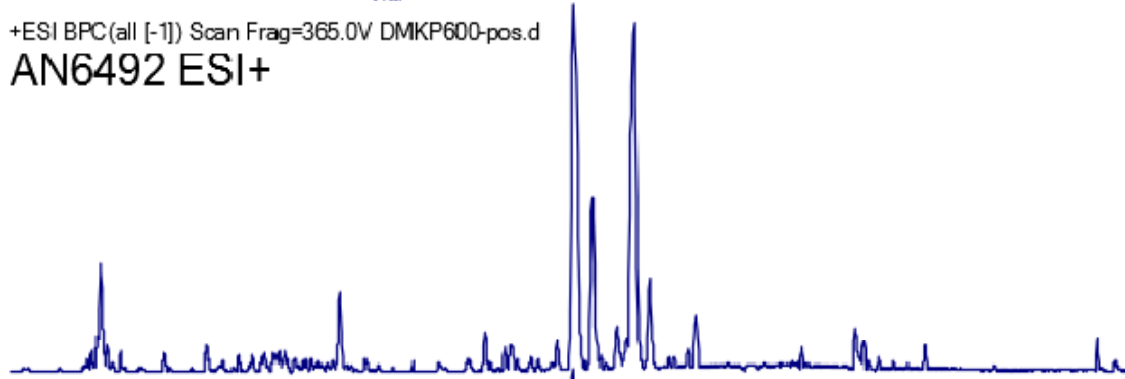
+ESI BPC(all [-1]) Scan Frag=365.0V DMKP599-pos.d

AN6420 ESI+



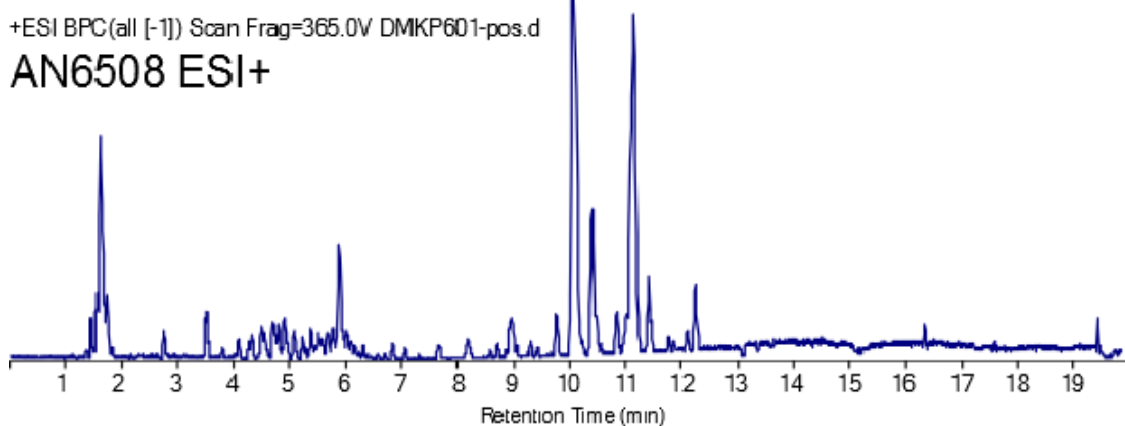
+ESI BPC(all [-1]) Scan Frag=365.0V DMKP600-pos.d

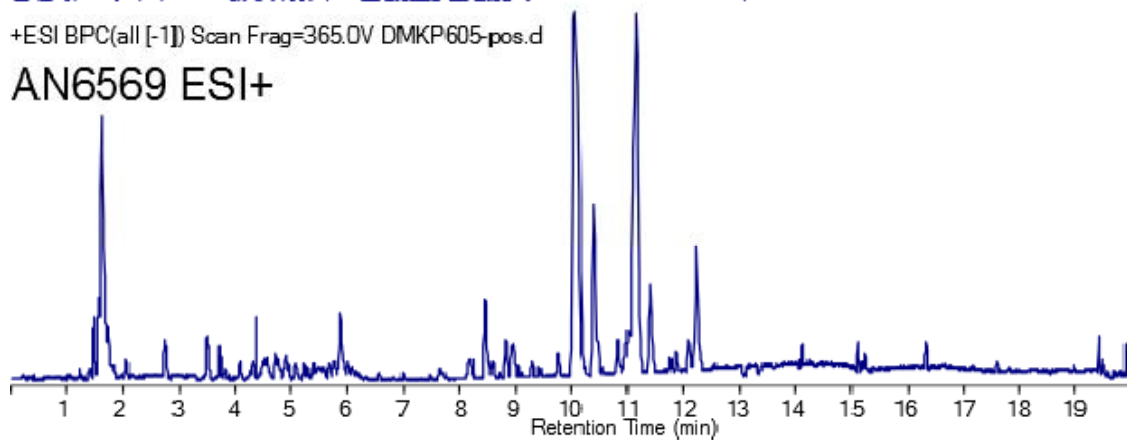
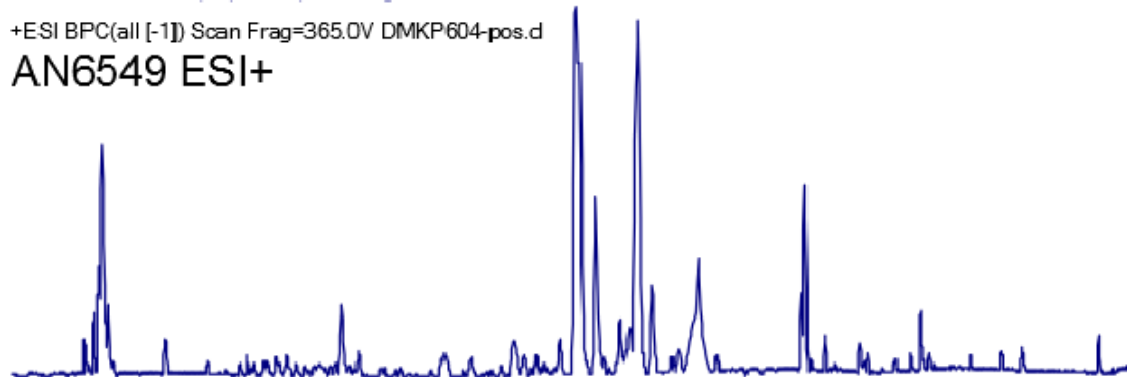
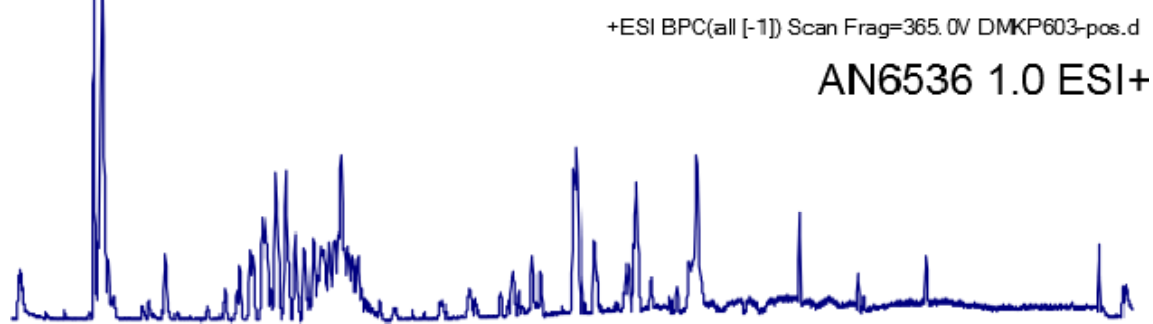
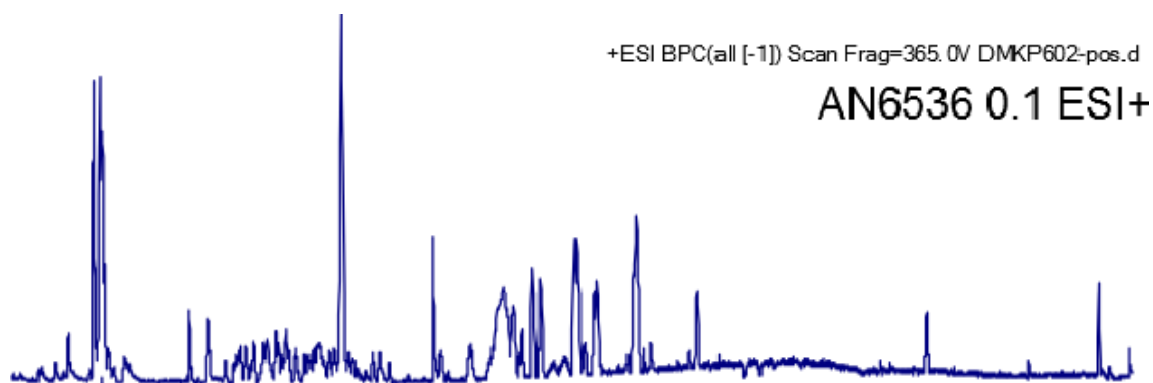
AN6492 ESI+



+ESI BPC(all [-1]) Scan Frag=365.0V DMKP601-pos.d

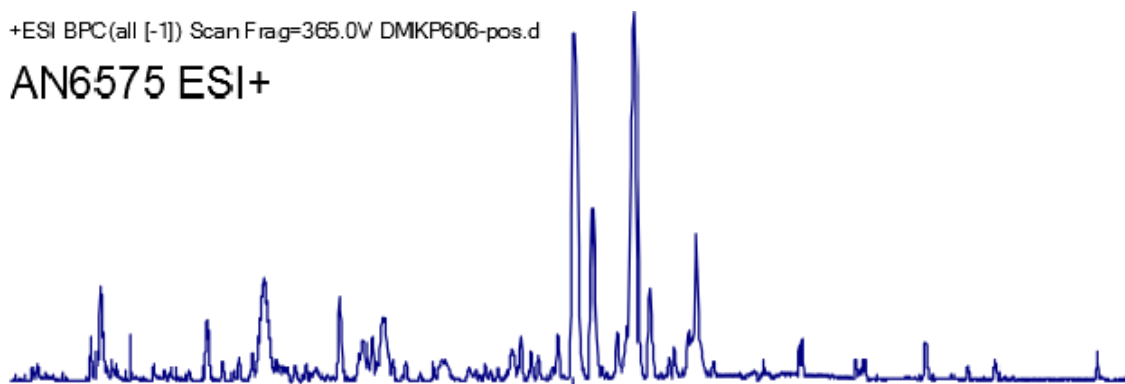
AN6508 ESI+





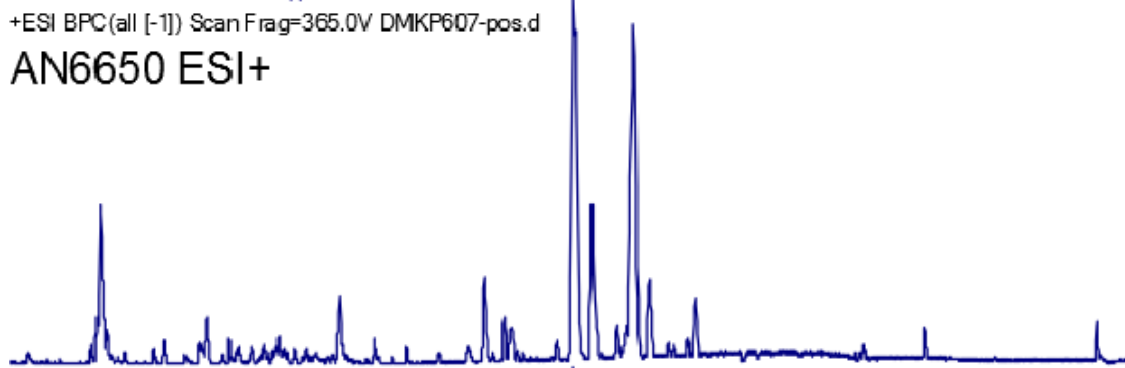
+ESI BPC(all [-1]) Scan Frag=365.0V DMKP606-pos.d

AN6575 ESI+



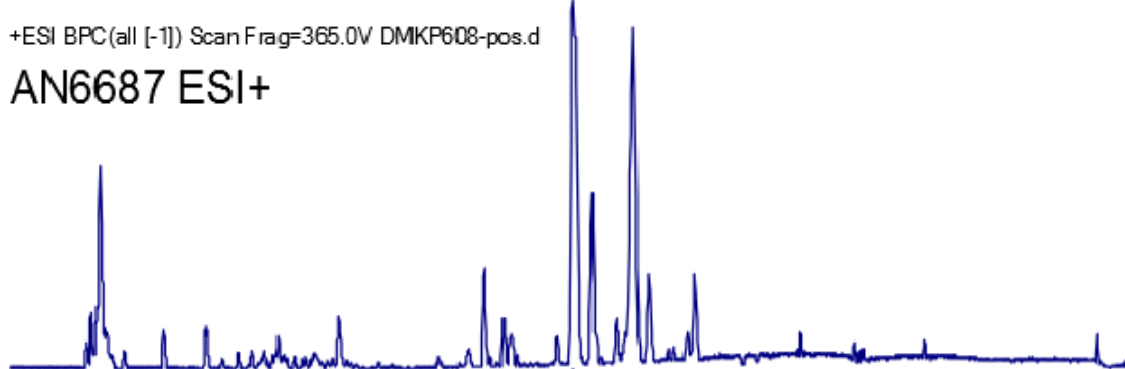
+ESI BPC(all [-1]) Scan Frag=365.0V DMKP607-pos.d

AN6650 ESI+



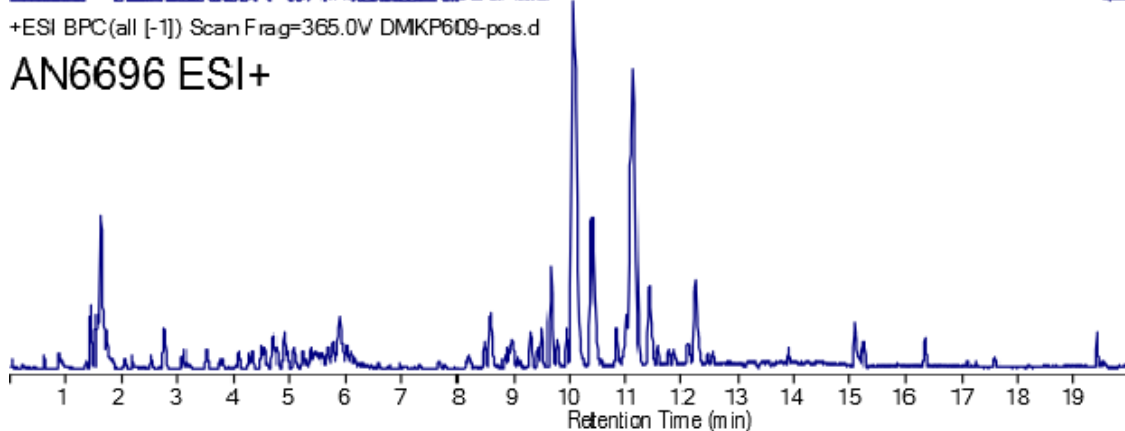
+ESI BPC(all [-1]) Scan Frag=365.0V DMKP608-pos.d

AN6687 ESI+



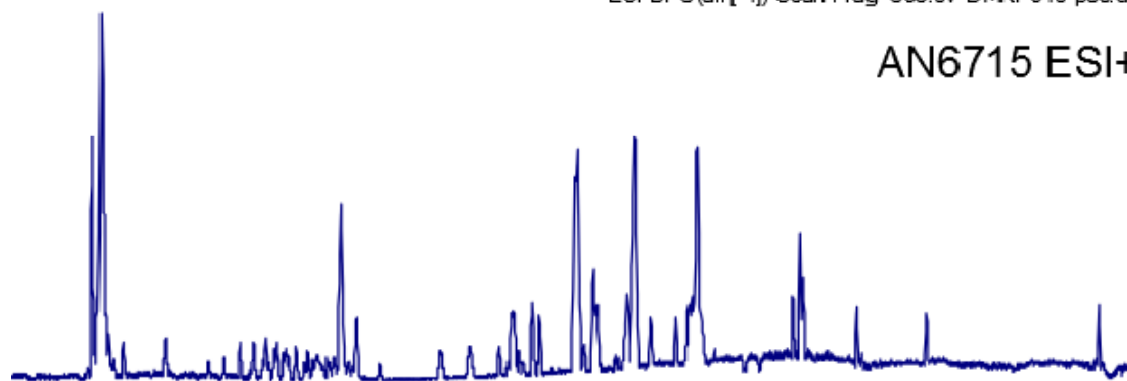
+ESI BPC(all [-1]) Scan Frag=365.0V DMKP609-pos.d

AN6696 ESI+



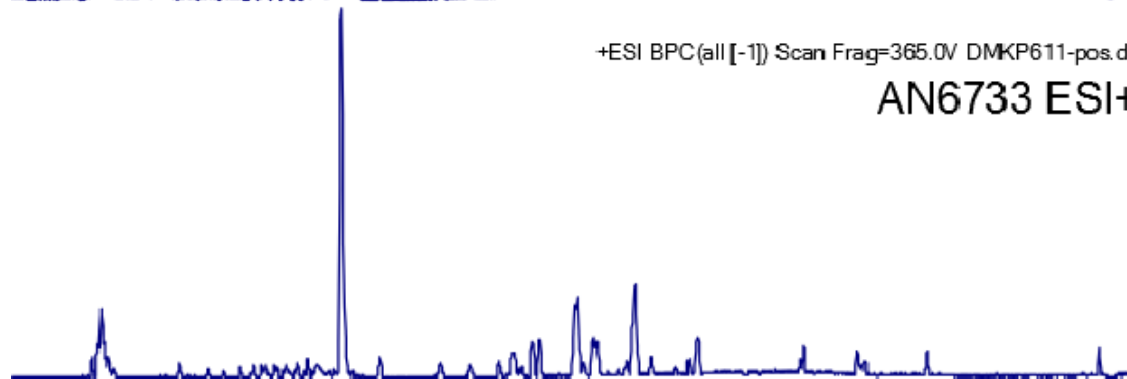
+ESI BPC(all [-1]) Scan Frag=365.0V DMKP610-pos.d

AN6715 ESI+



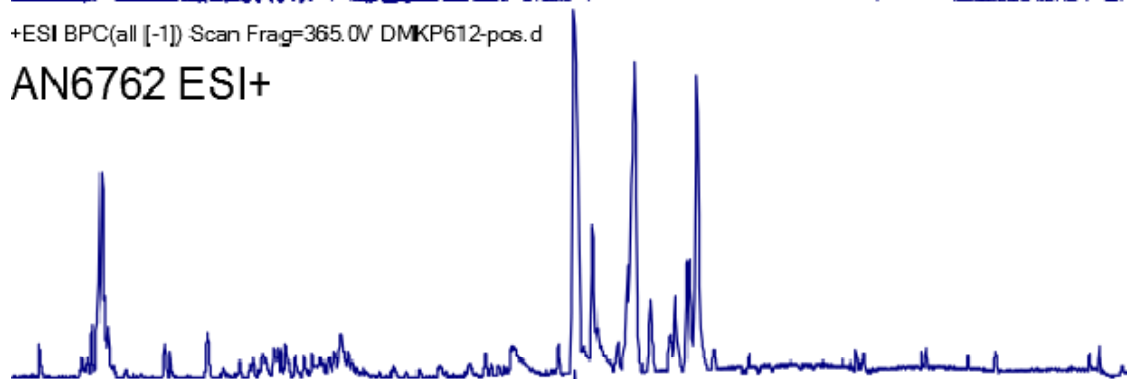
+ESI BPC(all [-1]) Scan Frag=365.0V DMKP611-pos.d

AN6733 ESI+



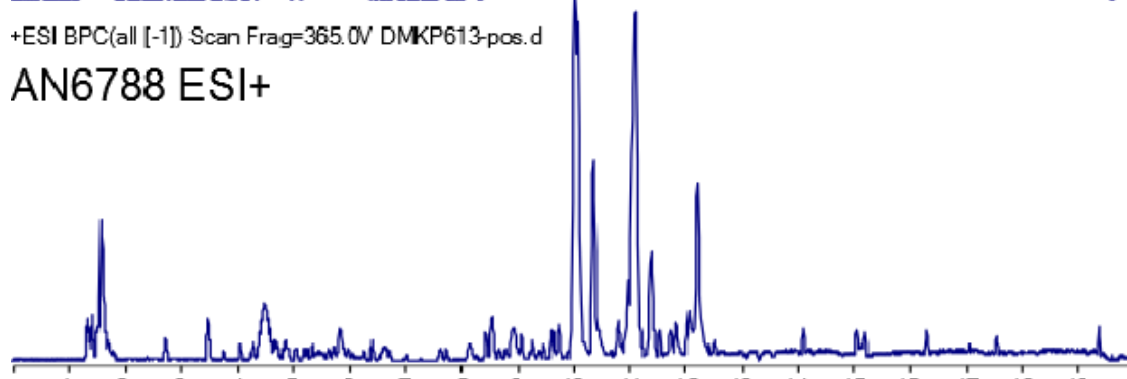
+ESI BPC(all [-1]) Scan Frag=365.0V DMKP612-pos.d

AN6762 ESI+



+ESI BPC(all [-1]) Scan Frag=365.0V DMKP613-pos.d

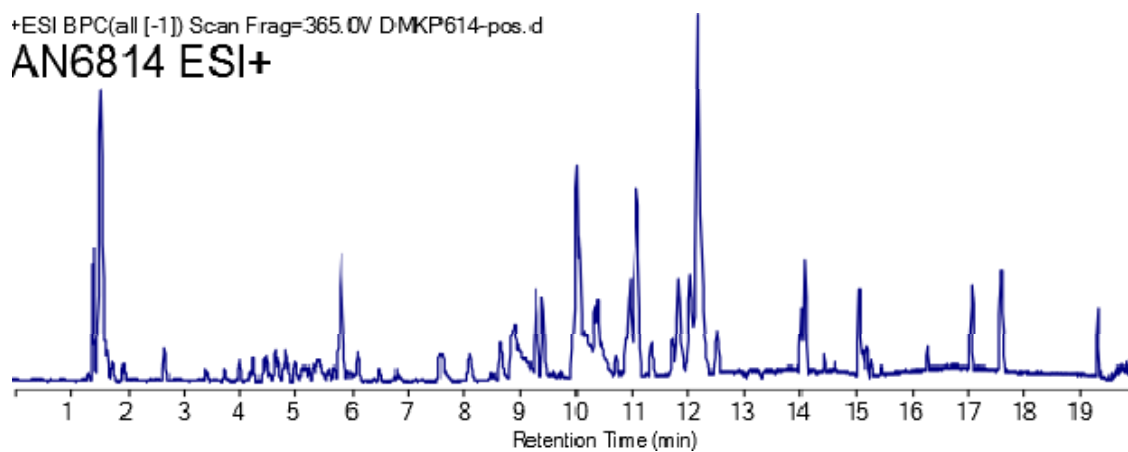
AN6788 ESI+



Retention Time (min)

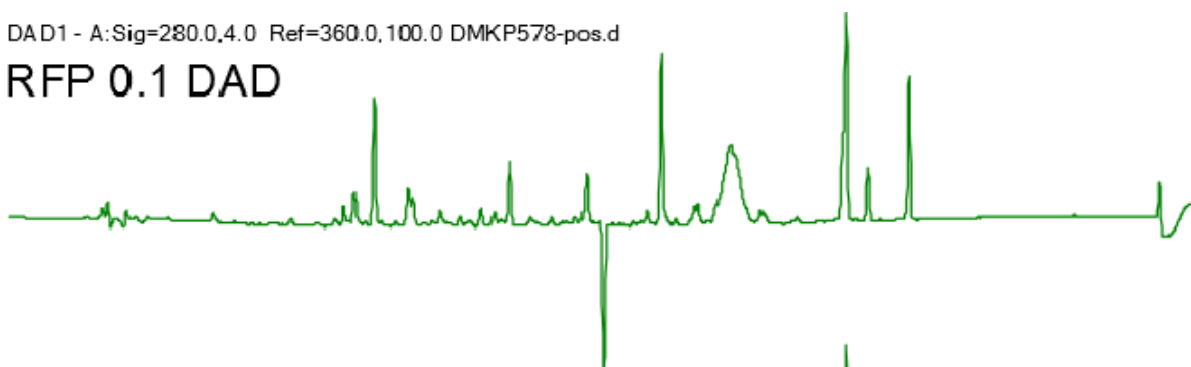
+ESI BPC(all [-1]) Scan Frag=365.0V DMKP614-pos.d

AN6814 ESI+



DAD1 - A:Sig=280.0,4.0 Ref=360.0,100.0 DMKP578-pos.d

RFP 0.1 DAD



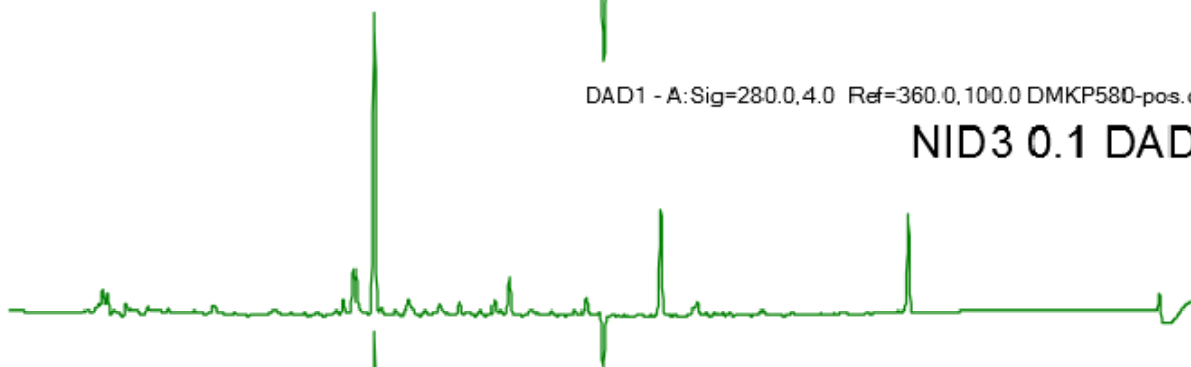
DAD1 - A:Sig=280.0,4.0 Ref=360.0,100.0 DMKP579-pos.d

RFP 1.0 DAD



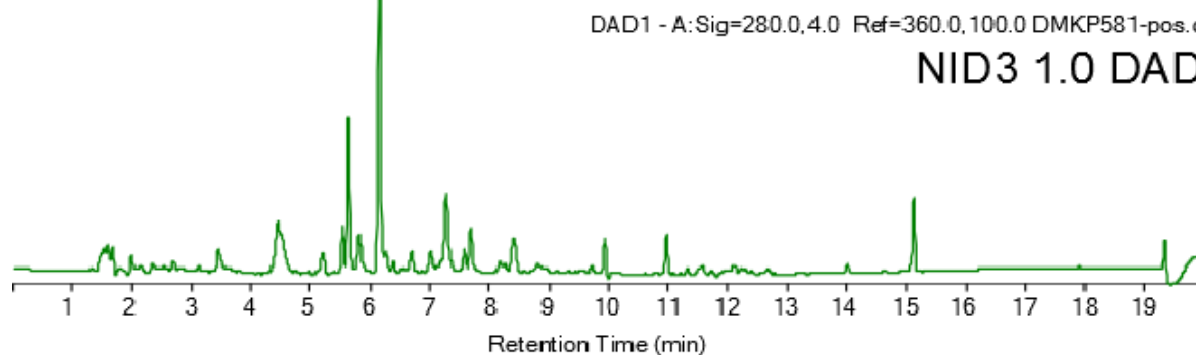
DAD1 - A:Sig=280.0,4.0 Ref=360.0,100.0 DMKP580-pos.d

NID3 0.1 DAD



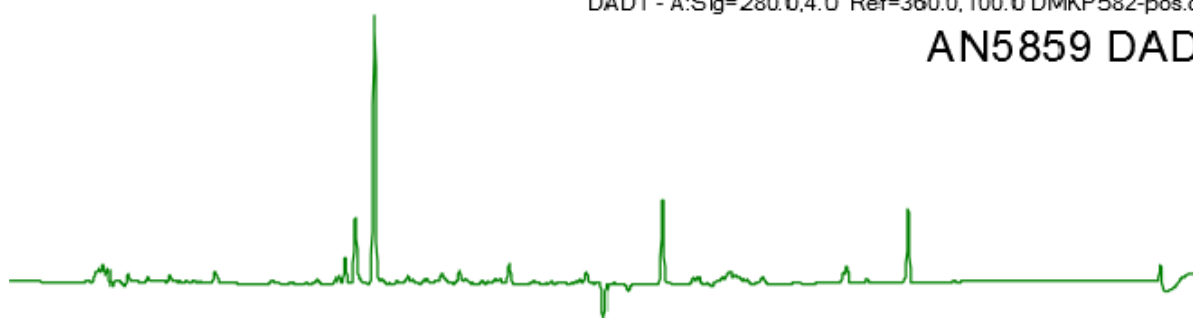
DAD1 - A:Sig=280.0,4.0 Ref=360.0,100.0 DMKP581-pos.d

NID3 1.0 DAD



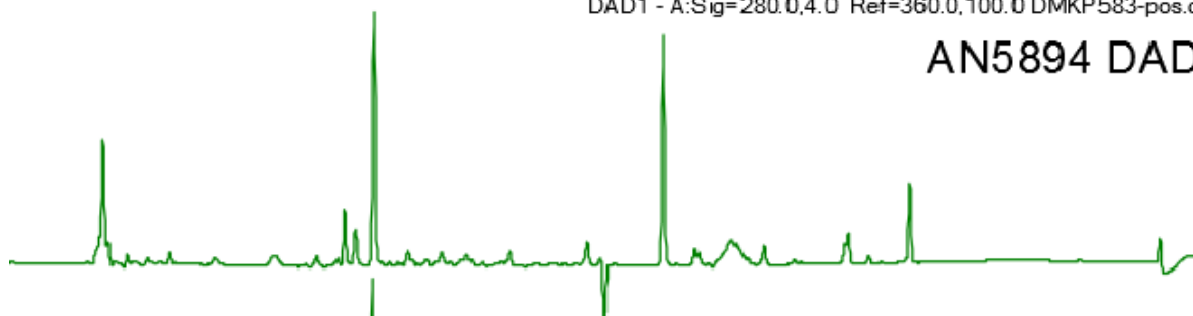
DAD1 - A:Sig=280.0,4.0 Ref=360.0,100.0 DMKP582-pos.d

AN5859 DAD



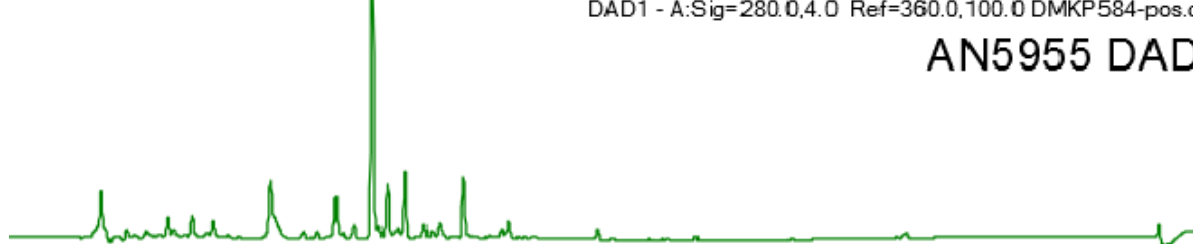
DAD1 - A:Sig=280.0,4.0 Ref=360.0,100.0 DMKP583-pos.d

AN5894 DAD



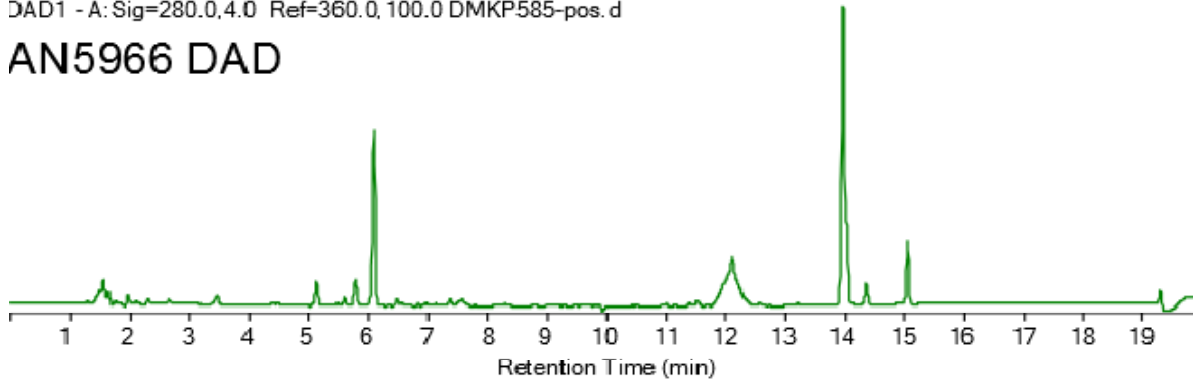
DAD1 - A:Sig=280.0,4.0 Ref=360.0,100.0 DMKP584-pos.d

AN5955 DAD



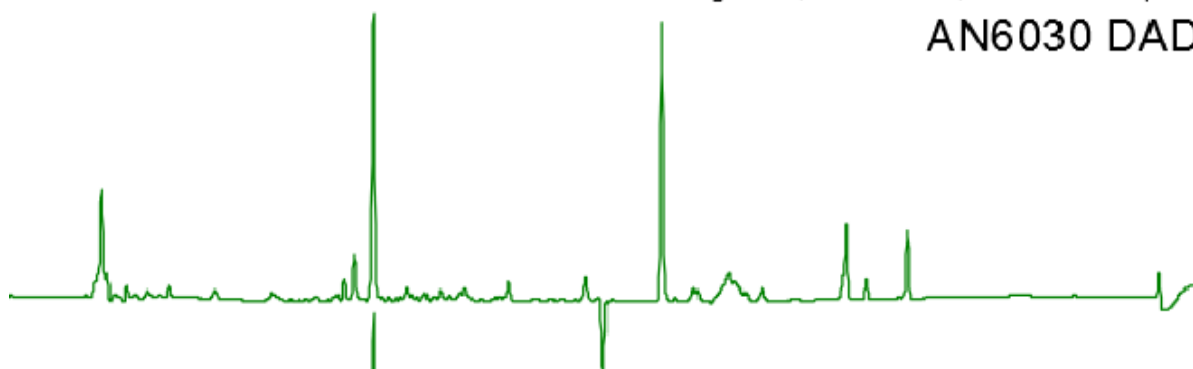
DAD1 - A:Sig=280.0,4.0 Ref=360.0,100.0 DMKP585-pos.d

AN5966 DAD



DAD1 - A:Sig=280.0,4.0 Ref=360.0,100.0 DMKP586-pos.d

AN6030 DAD



DAD1 - A:Sig=280.0,4.0 Ref=360.0,100.0 DMKP587-pos.d

AN6062 DAD



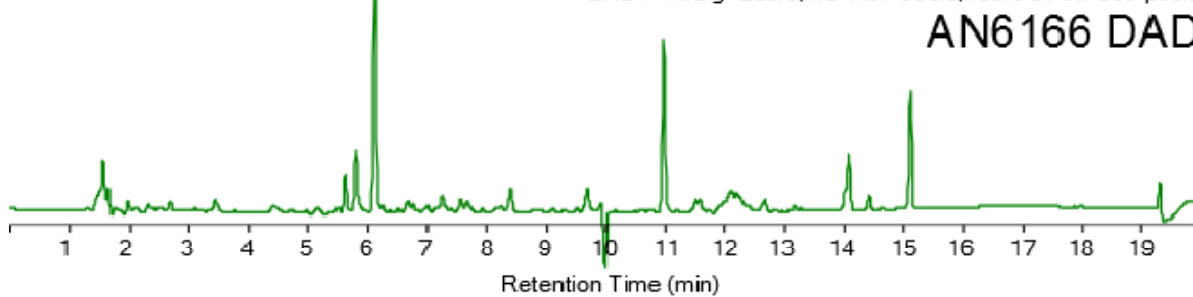
DAD1 - A:Sig=280.0,4.0 Ref=360.0,100.0 DMKP588-pos.d

AN6091 DAD



DAD1 - A:Sig=280.0,4.0 Ref=360.0,100.0 DMKP589-pos.d

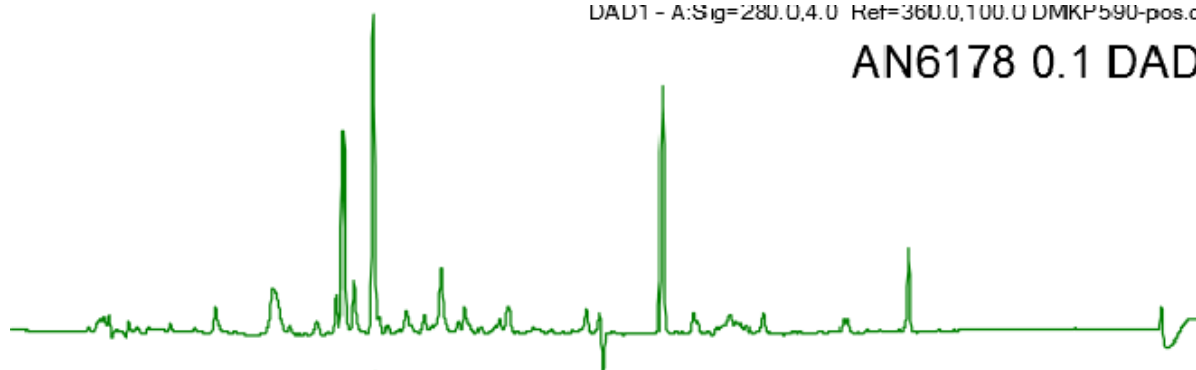
AN6166 DAD



Retention Time (min)

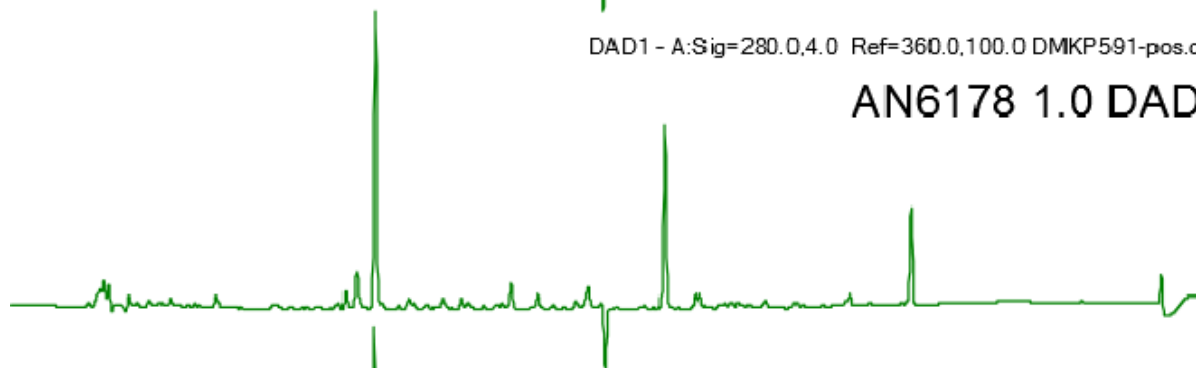
DAD1 - A:Sig=280.0,4.0 Ref=360.0,100.0 DMKP590-pos.d

AN6178 0.1 DAD



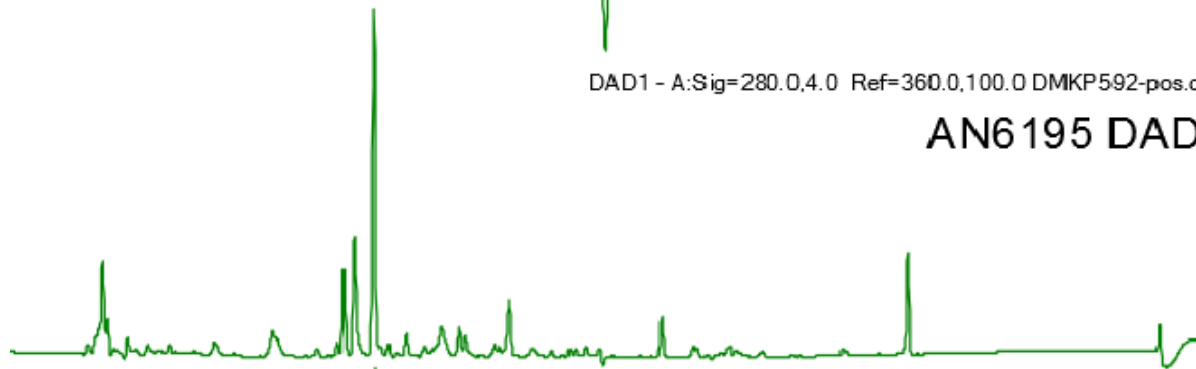
DAD1 - A:Sig=280.0,4.0 Ref=360.0,100.0 DMKP591-pos.d

AN6178 1.0 DAD



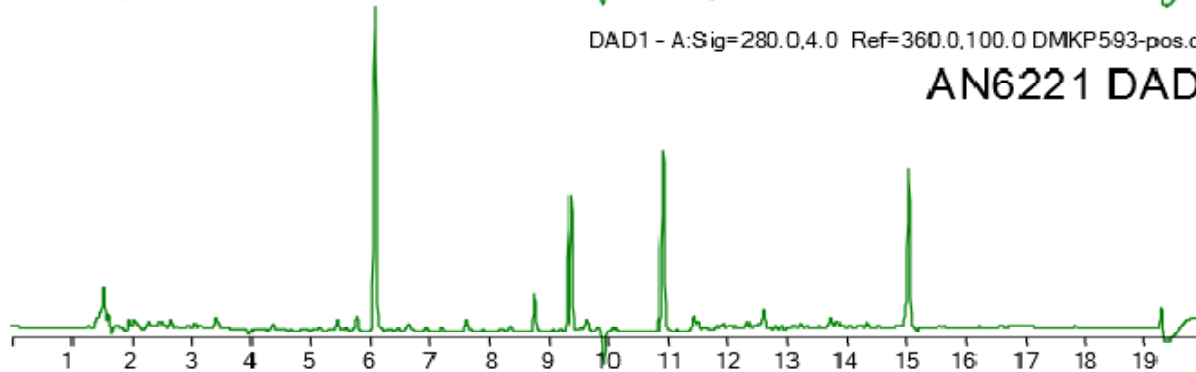
DAD1 - A:Sig=280.0,4.0 Ref=360.0,100.0 DMKP592-pos.d

AN6195 DAD



DAD1 - A:Sig=280.0,4.0 Ref=360.0,100.0 DMKP593-pos.d

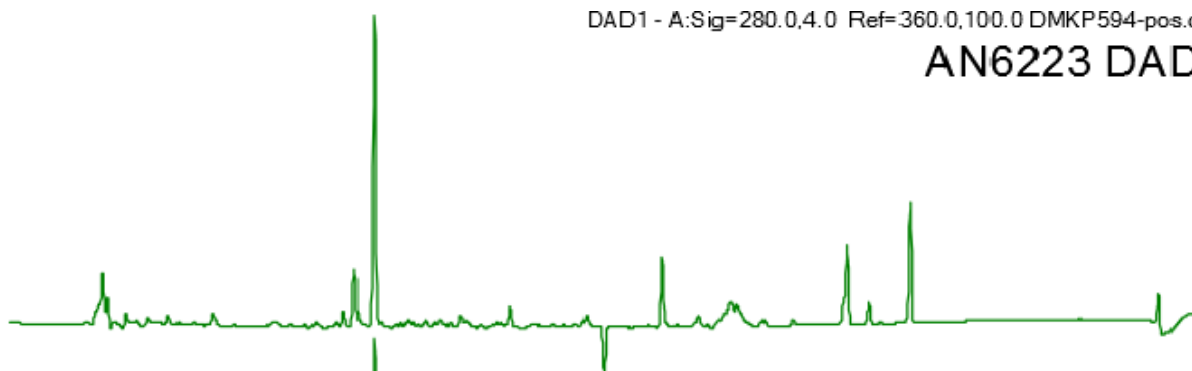
AN6221 DAD



Retention Time (min)

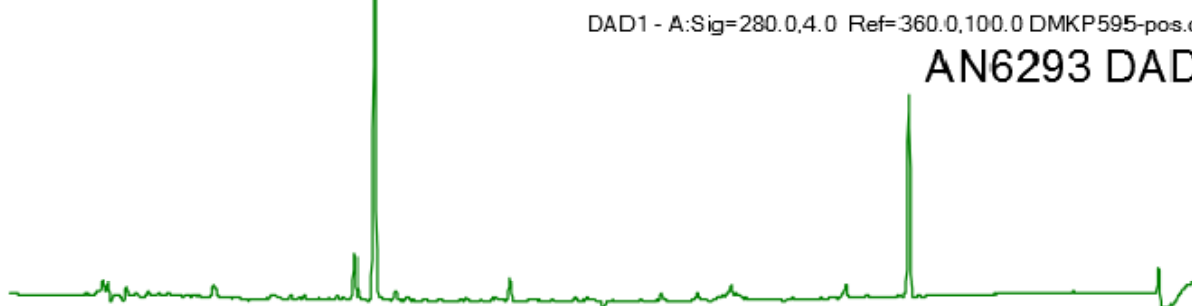
DAD1 - A:Sig=280.0,4.0 Ref=360.0,100.0 DMKP594-pos.d

AN6223 DAD



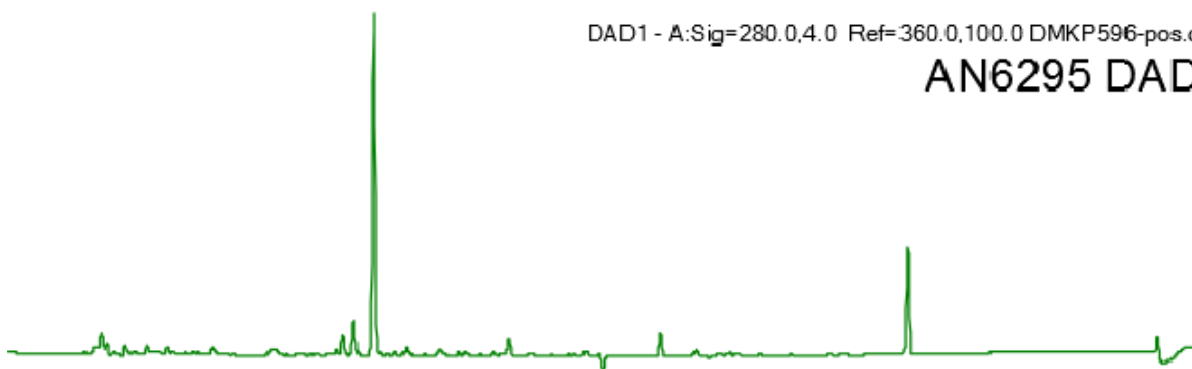
DAD1 - A:Sig=280.0,4.0 Ref=360.0,100.0 DMKP595-pos.d

AN6293 DAD



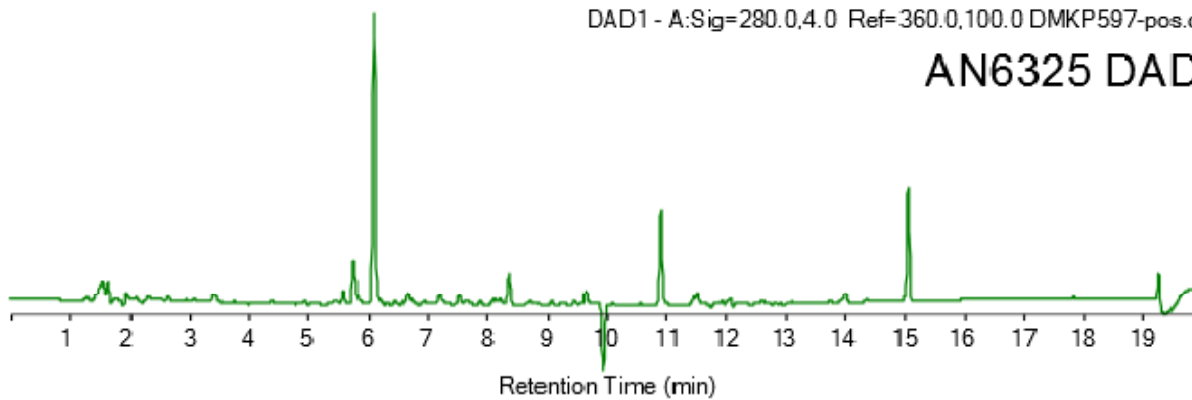
DAD1 - A:Sig=280.0,4.0 Ref=360.0,100.0 DMKP596-pos.d

AN6295 DAD



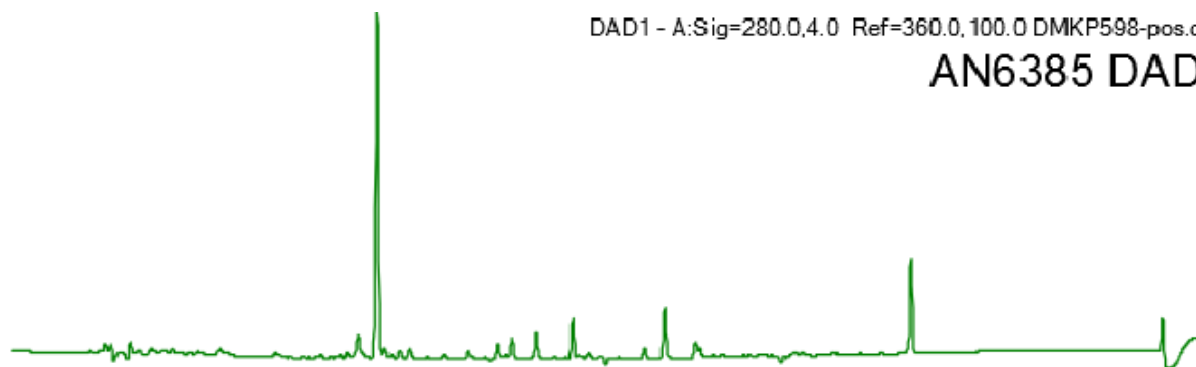
DAD1 - A:Sig=280.0,4.0 Ref=360.0,100.0 DMKP597-pos.d

AN6325 DAD



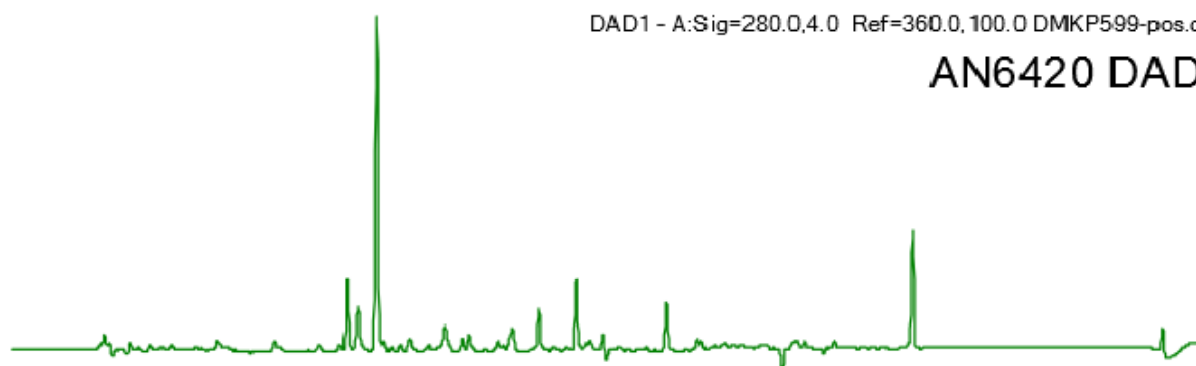
DAD1 - A:Sig=280.0,4.0 Ref=360.0,100.0 DMKP598-pos.d

AN6385 DAD



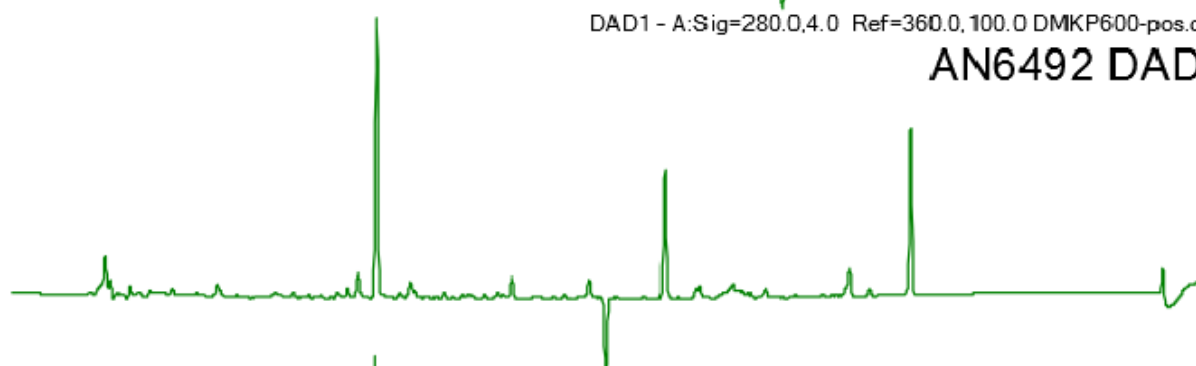
DAD1 - A:Sig=280.0,4.0 Ref=360.0,100.0 DMKP599-pos.d

AN6420 DAD



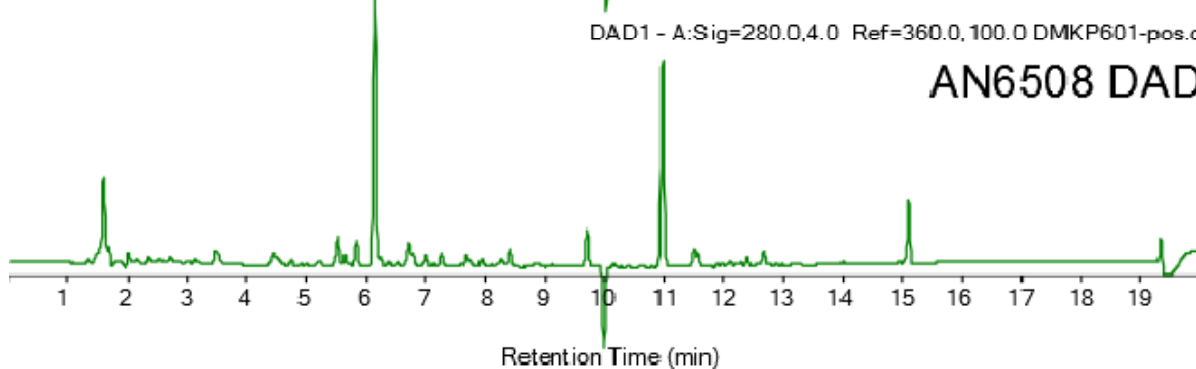
DAD1 - A:Sig=280.0,4.0 Ref=360.0,100.0 DMKP600-pos.d

AN6492 DAD



DAD1 - A:Sig=280.0,4.0 Ref=360.0,100.0 DMKP601-pos.d

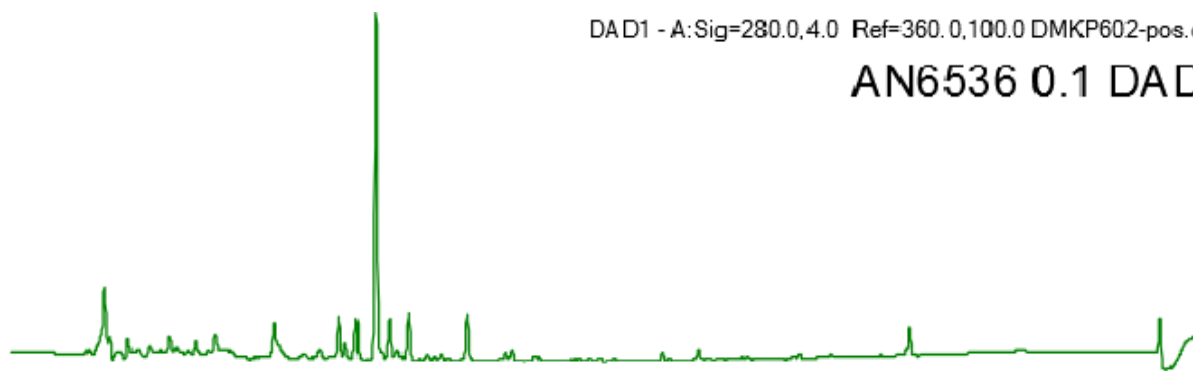
AN6508 DAD



Retention Time (min)

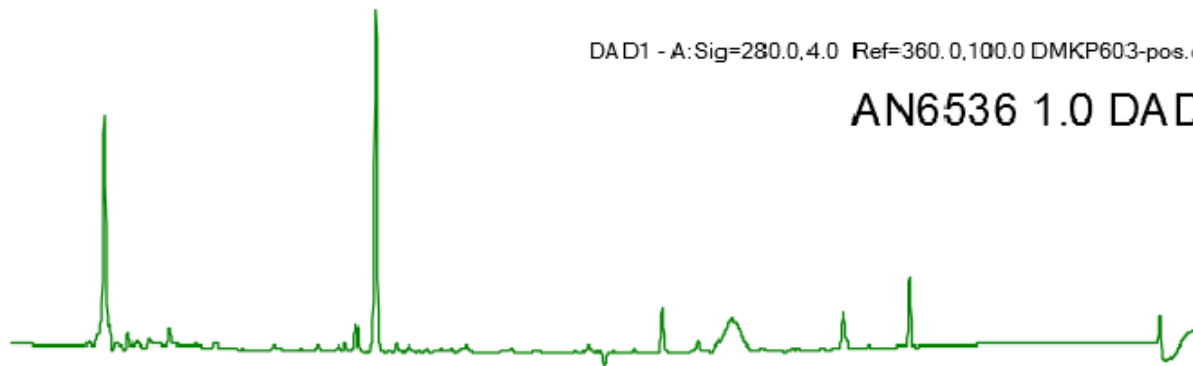
DAD1 - A:Sig=280.0,4.0 Ref=360.0,100.0 DMKP602-pos.d

AN6536 0.1 DAD



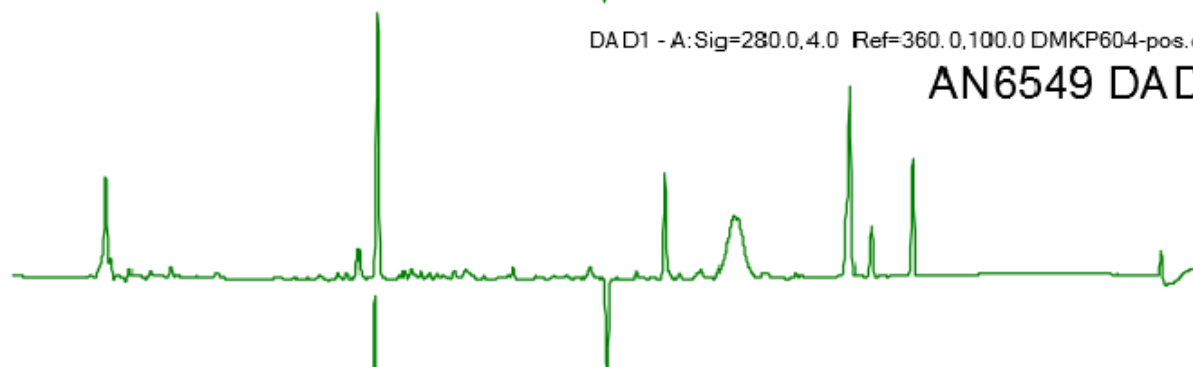
DAD1 - A:Sig=280.0,4.0 Ref=360.0,100.0 DMKP603-pos.d

AN6536 1.0 DAD



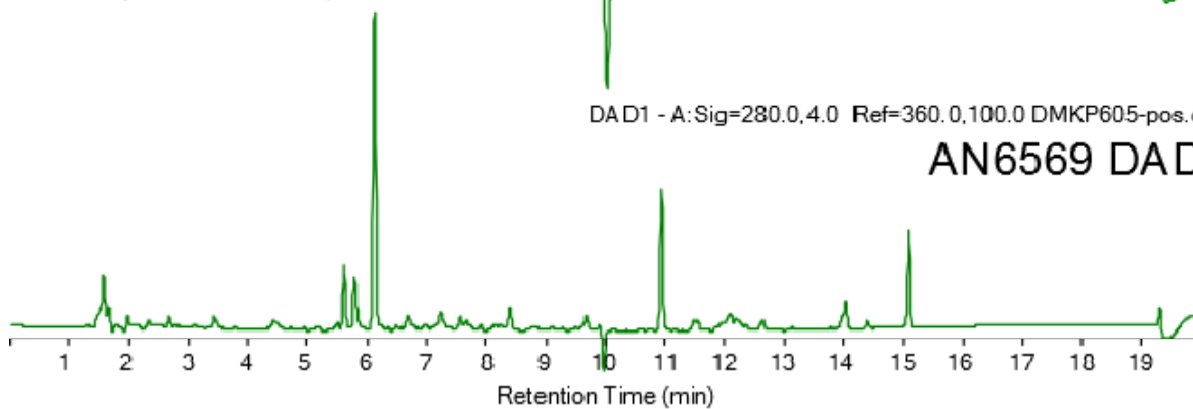
DAD1 - A:Sig=280.0,4.0 Ref=360.0,100.0 DMKP604-pos.d

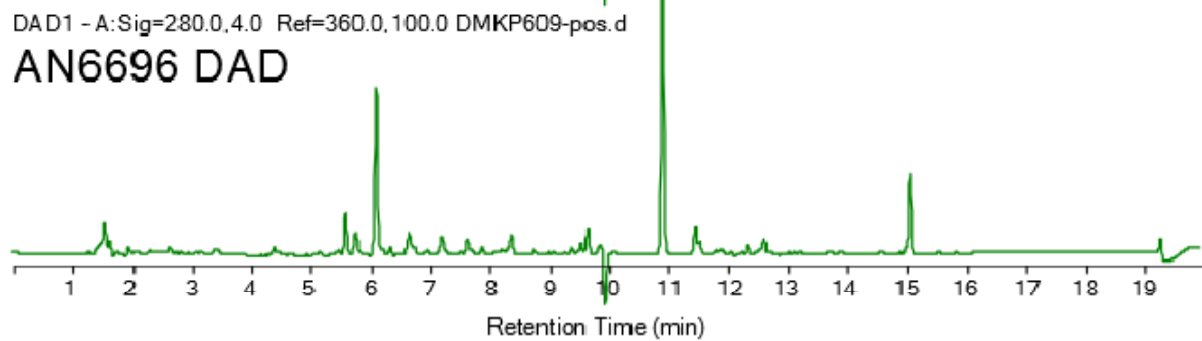
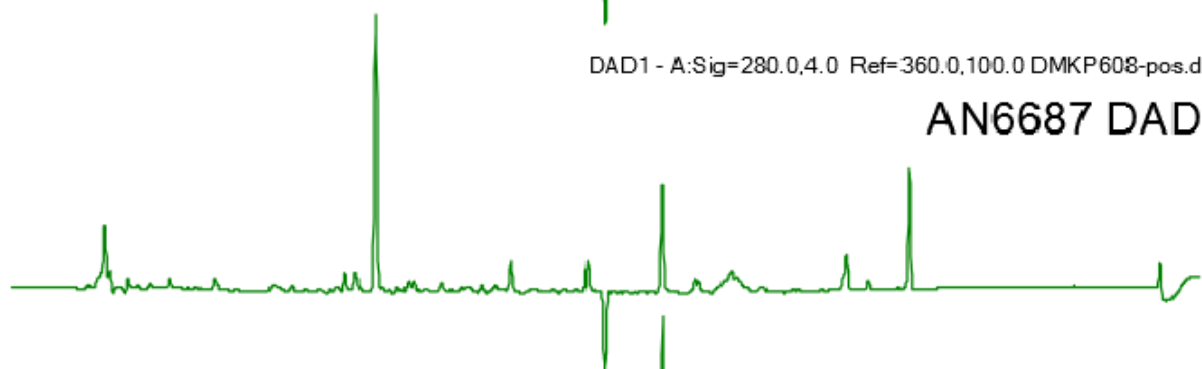
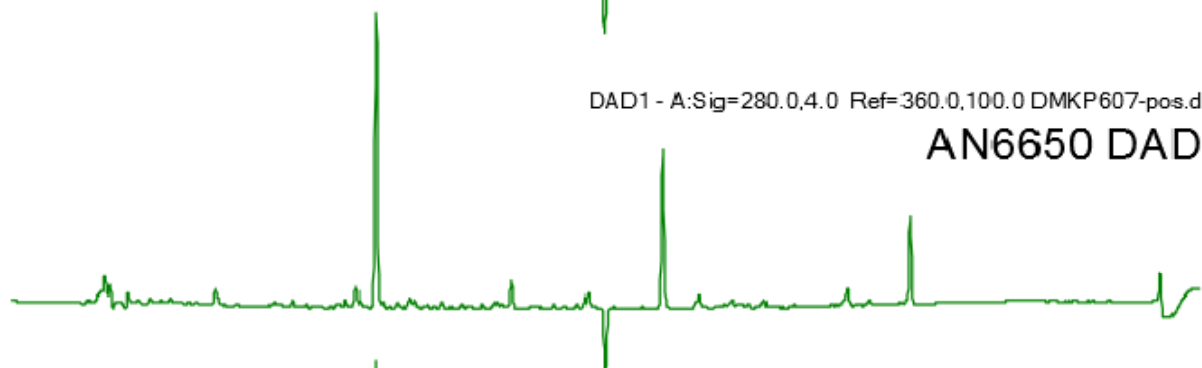
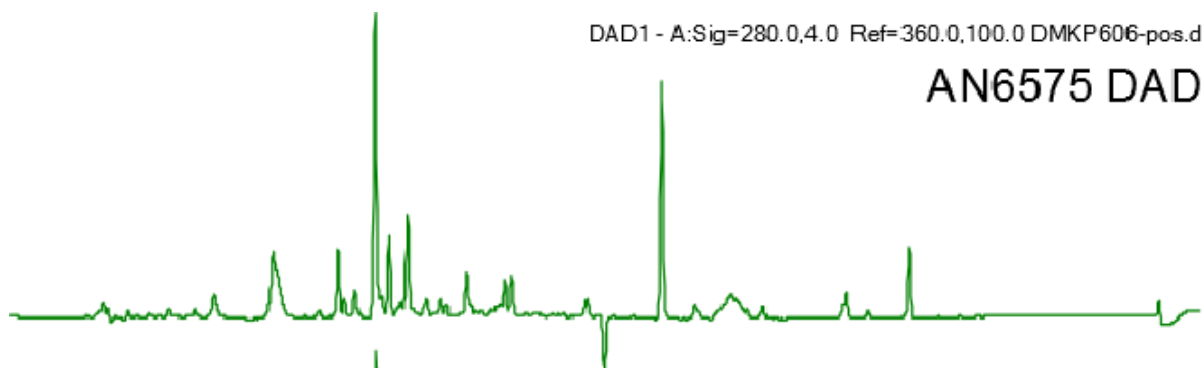
AN6549 DAD

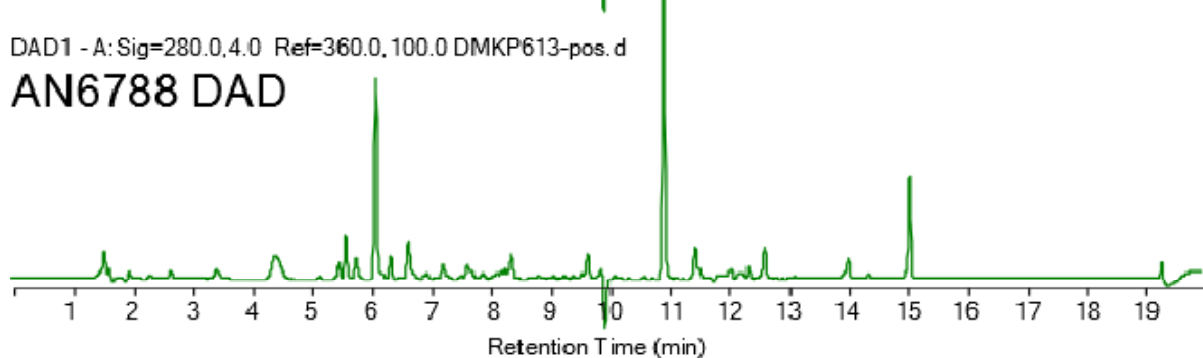
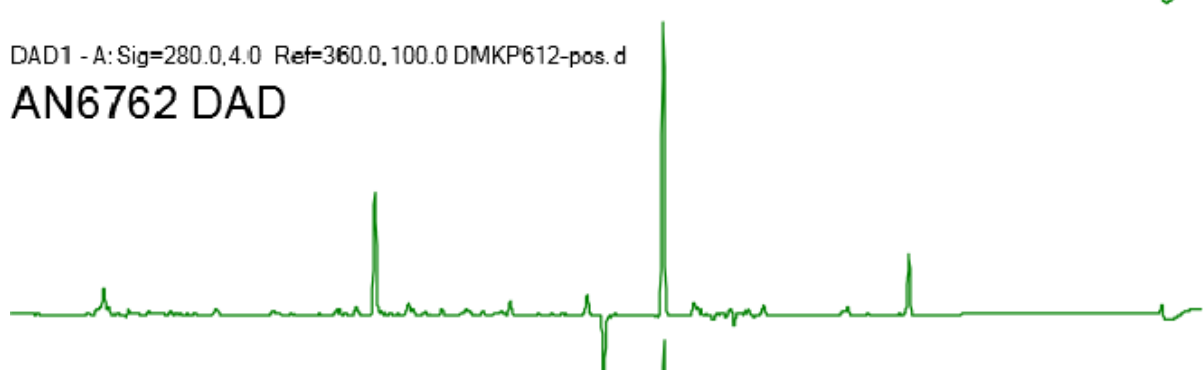
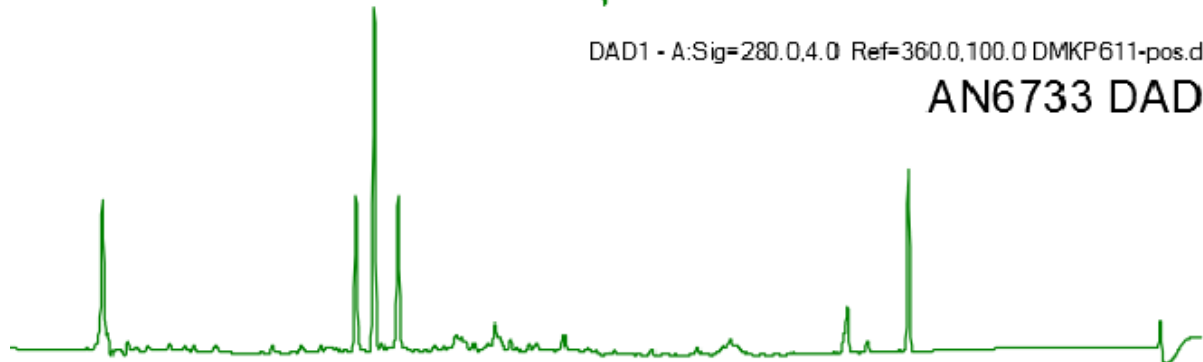
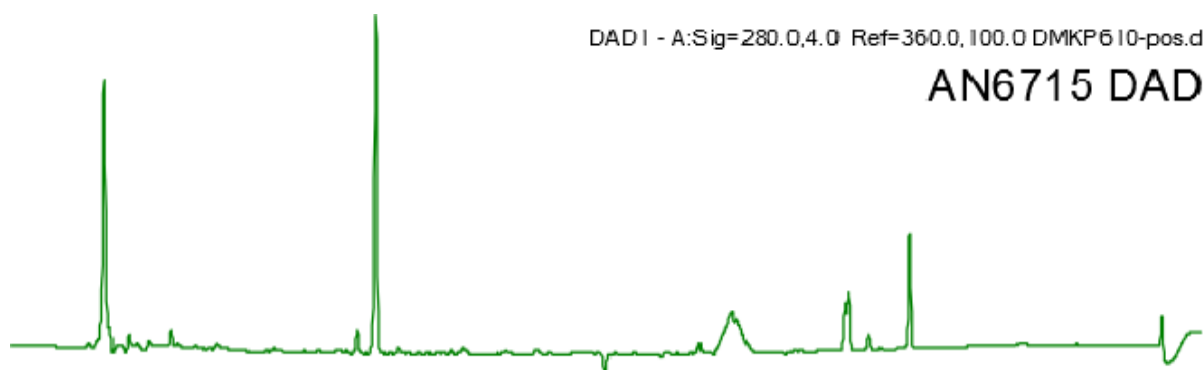


DAD1 - A:Sig=280.0,4.0 Ref=360.0,100.0 DMKP605-pos.d

AN6569 DAD

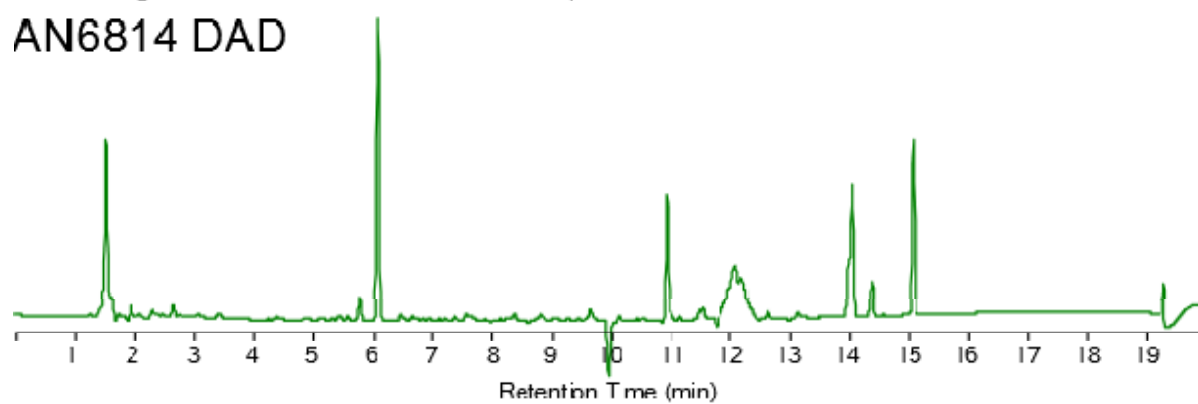






DAD1 - A: Sig=280.0,4.0 Ref=360.0,100.0 DMKP614-pos. d

AN6814 DAD



Appendix C

Supporting information for Chapter 7

Supplementary materials for:

PHUSER (Primer Help for USER): A novel tool for USER fusion primer design

Olsen *et al.* 2011

PROOF OF CONCEPT

MATERIALS AND METHODS

Sequence retrieval: The genome of *Aspergillus niger* strains ATCC1015 was recently sequenced and annotated (1). DNA sequences were retrieved from the DOE Joint Genome Institute database (<http://genome.jgi-psf.org/Aspni5/Aspni5.home.html>). The whole coding sequence including introns was amplified from start codon to stop codon, promoter and terminator were present in the cloning vector.

Primer design: The sequence of each gene was divided into PCR-friendly sizes ranging from 1.7 – 2.7 kb. The sequences were submitted to PHUSER in FASTA format, and the tab “1) Basic: Predefined fusion cassettes” was selected to choose the AsiSI/Nb.BtsI cassette, which was held by the target cloning vector, pU1111-1 (2).

PCR amplification: PCR fragments for USER fusion were produced by PCR using PfuX7 polymerase (3), and the primers were designed by PHUSER (table S1). Genomic DNA from *A. niger* strain ATCC1015 was purified with the FastDNA® SPIN Kit for Soil (Q-BIOgene). For amplification of genes from genomic DNA the reaction conditions were: initial denaturation (98°C; 30s) followed by 32 PCR cycles (98°C; 10s, T_m (lowest in primer-pair)+5°C; 30s, 72°C; 30s/kb DNA) and a final extension step (72°C; 5 min). PCR products were purified using the Illustra™ GFX PCR DNA and Gel Band Purification kit (GE Healthcare).

Cloning and transformation: The pU1111-1 vector was prepared for cloning by digestion with AsiSI and subsequent nicking with Nb.BtsI (both from New England Biolabs). 2 µg of pU1111-1 was digested in 40 U of AsiSI overnight at 37° C in 100 µL total volume. Subsequently, another 10 U of AsiSI and 10 U of Nb.BtsI was added and the mixture incubated at 37° C for 2 h. The digested vector was purified with Illustra™ GFX PCR DNA and Gel Band Purification kit (GE Healthcare). USER cloning was performed using 0.01 pmol of linearized pU1111-1 plasmid, equal amounts of insert fragments and 1 U USER™ enzyme mix (New England Biolabs) in 13 µL total volume. The mixture incubated for 20 min. at 37° C followed by 20

min. at 25° C. All of the cloning mixture was transformed into chemically competent *Escherichia coli* DH5α cells. Purified plasmid from the resulting clones were digested with NruI endonuclease and a clone was selected based on the predicted band pattern; a pattern distinctive of PKS insertion. The plasmid originating from the selected clone was sequenced (StarSeq, Germany). The CLC software was used for all computational handling of sequence.

TABLES

Table S1. Primers designed by PHUSER.

Name	Sequence 5'-3'	Product size
Gene ID: 44965		
FW-44965-1	AGAGCGAUATGCCAGGCCTTGTACAC	2624
RV-44965-1	ACGGGGTCCUCGGTACATTCCCCCATAGTTT	
FW-44965-2	AGGACCCCGUTCTCGTCACCCCTACCAT	2740
RV-44965-2	TCTGCGAUTTAAGCATCCAGCTCCTTTGT	
Gene ID: 55896		
FW-55896-1	AGAGCGAUATGGAGGGTCCATCTCGTGTGT	2250
RV-55896-1	AGAACCGAUATGAAGCGATTGACCGGGAC	
FW-55896-2	ATCGGTTTCUGCCAAGGCAAACATTGGAC	2170
RV-55896-2	AAAGAAUCGGACAGTGCAAGTTGCG	
FW-55896-3	ATTCTTUGACTGCGCTGCTGCGGAG	2260
RV-55896-3	TCTGCGAUTTAATTGGCCATGGCGTTGCCA	
Gene ID: 128601		
FW-128601-1	AGAGCGAUATGTCTTCTTCACCCAAAGAGCCC	2040
RV-128601-1	AGTGGCGCUAAAAGAAGCCGGCGGCGTA	
FW-128601-2	AGCGCCACUGATGCCATCCGCATTGCCTTCT	2240
RV-128601-2	AGTGTUCCAGGAGAGGCTTGGTAGCG	
FW-128601-3	AGACACUCGGTGCCTTGTACTCCGCCTACAC	2310
RV-128601-3	ACCGTCGAUCTTGGGGCCAGAGCCTCGTTCA	
FW-128601-4	ATCGACGGUACCATTCTACCTGGACGAACTGTTT	2730
RV-128601-4	ACCGCAAGUGGCCGCAAGACGGGGCAG	
FW-128601-5	ACTTGCGGUGTTGTGAGCCCAGCTCC	2537
RV-128601-15	TCTGCGAUCCGGCAATCTTGGCCTTGC	
Gene ID: 194381		
FW-194381-1	AGAGCGAUATGGCAATCATGGAACAG	2259
RV-194381-1	ACCGGCGCUGTTATTGACGTTCCGTACT	
FW-194381-2	AGCGCCGGUGTCCAAGTCTCAGCAATCCCTCTA	2450
RV-194381-2	ATCTTGCAUGATTCCGGCAGCGGCCAC	
FW-194381-3	ATGCAAGAUTCTGTATACGATCCTCTA	2255
RV-194381-3	TCTGCGAUTTATTATCAAGCGGAGCT	
Gene ID: 211885		
FW-211885-1	AGAGCGAUATGAGTCCCAGCAAGAAT	1776
RV-211885-1	ACGCCGAGCAAUGAAGCAGTTCCATGCAGT	
FW-211885-2	ATTGCTCGGCGUTTTCACCGGCCAGGGAGC	2310
RV-211885-2	ATAGCCGCUCAGCATTCTTGAGCAAAGCCTGT	
FW-211885-3	AGCGGCTAUTCCATACTGTCAATGACGG	1820
RV-211885-3	AGCATTUGTAATACCCACTGGACTGT	
FW-211885-4	AAATGCUCCACCCACGGGGACTGT	2112
RV-211885-4	TCTGCGAUTCAGACCACCTCCTGTGTACTGAT	
Gene ID: 225574		
FW-225574-1	AGAGCGAUATGGATAGTCAAGACTCGTCTCCC	1620
RV-225574-1	AGCTTTCTUTTAGGGATGCCGGGTGTGC	
FW-225574-2	AAGAAAGCUTAGTCGAGCATGAACGATACATCA	2310
RV-225574-2	AGAATCAAUATGCGAGCTATCGCTGGC	
FW-225574-3	ATTGATTCTTGTGAGCTCGAGCTCGTC	2030
RV-225574-3	ATCGTGGCCGCUTCTTCGTATGAGGAAACGATG	
FW-225574-4	AGCGGCCACGAUGCCATGTGTTTACGCTACT	2027
RV-225574-4	TCTGCGAUTCACGGGGCCTTCATTTC	

..

REFERENCES

1. Andersen, M., Salazar, M., Schaap, P., Vondervoort, P.v.d., Culley, D., Thykaer, J., Frisvad, J., Nielsen, K., Albang, R., Albermann, K. *et al.* (2011) Comparative genomics of citric-acid producing *Aspergillus niger* ATCC 1015 versus enzyme-producing CBS 513.88. *Genome Research*, **Accepted for publication**.
2. Hansen, B.G., Salomonsen, B., Nielsen, M.T., Nielsen, J.B., Hansen, N.B., Nielsen, K.F., Regueira, T.B., Nielsen, J., Patil, K.R. and Mortensen, U.H. (2011) A versatile gene expression and characterization system for *Aspergillus*: heterologous expression of the gene encoding the polyketide synthase from the mycophenolic acid gene cluster from *Penicillium brevicompactum* as a case study. *Appl Environ Microbiol.*
3. Norholm, M.H. (2010) A mutant Pfu DNA polymerase designed for advanced uracil-excision DNA engineering. *BMC Biotechnol*, **10**, 21.

Appendix D

Supporting information for Chapter 9

Table A1. Fungal strains used and constructed in this study. Strains in red are intended but have not yet been constructed.

Name	Organism	Genotype
IBT 29539 [Nielsen <i>et al.</i> , 2006]	<i>A. nidulans</i>	<i>argB2, pyrG89, veA1, nkuAΔ</i>
IS1-PKS44	<i>A. nidulans</i>	<i>argB2, pyrG89, veA1, nkuAΔ IS1-PKS44</i>
IS1-Acar44	<i>A. nidulans</i>	<i>argB2, pyrG89, veA1, nkuAΔ IS1-Acar44</i>
IS1- <i>albA</i>	<i>A. nidulans</i>	<i>argB2, pyrG89, veA1, nkuAΔ IS1-<i>albA</i></i>
IS1- <i>wA</i>	<i>A. nidulans</i>	<i>argB2, pyrG89, veA1, nkuAΔ IS1-<i>wA</i></i>
<i>yAΔ</i>	<i>A. nidulans</i>	<i>argB2, pyrG89, veA1, nkuAΔ yAΔ</i>
<i>wAΔ yAΔ</i>	<i>A. nidulans</i>	<i>argB2, pyrG89, veA1, nkuAΔ wAΔ yAΔ</i>
IS1-PKS44 <i>w</i>	<i>A. nidulans</i>	<i>argB2, pyrG89, veA1, nkuAΔ wAΔ IS1-PKS44</i>
IS1-Acar44 <i>w</i>	<i>A. nidulans</i>	<i>argB2, pyrG89, veA1, nkuAΔ wAΔ IS1-Acar44</i>
IS1- <i>albA</i> <i>w</i>	<i>A. nidulans</i>	<i>argB2, pyrG89, veA1, nkuAΔ wAΔ IS1-<i>albA</i></i>
IS1-PKS44 <i>wy</i>	<i>A. nidulans</i>	<i>argB2, pyrG89, veA1, nkuAΔ wAΔ yAΔ IS1-PKS44</i>
IS1-Acar44 <i>wy</i>	<i>A. nidulans</i>	<i>argB2, pyrG89, veA1, nkuAΔ wAΔ yAΔ IS1-Acar44</i>
IS1- <i>albA</i> <i>wy</i>	<i>A. nidulans</i>	<i>argB2, pyrG89, veA1, nkuAΔ wAΔ yAΔ IS1-<i>albA</i></i>
KB1001 [Chiang <i>et al.</i> , 2011]	<i>A. niger</i>	<i>pyrGΔ kusA::AFpyrG</i>
<i>napAΔ</i>	<i>A. niger</i>	<i>pyrGΔ kusA::AFpyrG napAΔ</i>
<i>albAΔ-napA</i>	<i>A. niger</i>	<i>pyrGΔ kusA::AFpyrG albAΔ::napA::hph</i>
<i>albAΔ-rfp</i>	<i>A. niger</i>	<i>pyrGΔ kusA::AFpyrG albAΔ::rfp::hph</i>
IS-NIG1- <i>napA</i>	<i>A. niger</i>	<i>pyrGΔ kusA::AFpyrG IS-NIG1::napA::hph</i>
<i>napΔ</i> (cluster)	<i>A. niger</i>	<i>pyrGΔ kusA::AFpyrG napΔ</i>
<i>albAΔ napAΔ</i>	<i>A. niger</i>	<i>pyrGΔ kusA::AFpyrG albAΔ napAΔ</i>

Table A2. Primers used in the study. See details in experimental section.

#	Primer name	Sequence 5'→3'
1	ori coli FW	ATCCCCACUACCGCATTAAGACCTCAGCG
2	ampR_RV	AGCTGCTUCGTCGATTAAACCCTCAGCG
3	p71_prom- ter_short_usF	AGCCCAATAUTAAGCTCCCTAATTGGCCC
4	p71_prom-ter_dsR	ATTACCTAGUGGGCGCTTACACAGTACA
5	ASN-IS1-US-Fw	AAGCAGCUACGCTTGGATAAGAAGTGAGAAGT
6	ASN-IS1-US-Rv	ATATTGGGCUTCTACACCGATATGGAGCTCTTT
7	ASN-IS1-DS-Fw	ACTAGGTAAUAAGTAAGCGTGCGTGAAACAG
8	ASN-IS1-DS-Rv	AGTGGGGAUTAATTATAACTGGTCAGCCACTGAC
9	Chk-IS1-US-Fw	GGAGTTTCGGTGGAGTCTACGGAG
10	Chk-IS1-DS-Rv	TGGGTACATACCTCTTCTGGGT
11	AlbA_US_FW	AAGCAGCUATCCGTCGGTCAGGCTGCT
12	AlbA_US_RV	ATATTGGGCUCACACGAGATGGACCCTCCAT
13	AlbA_DS_FW	ACTAGGTAAUTCCCGCTAGGACTGACAACA
14	AlbA_DS_RV	AGTGGGGAUCTCCCTCGTATTCGCCTCCT
15	Chk-IS2-US-Fw	ACTCAGCATGTTGGTTATATATTCGAG
16	Chk-IS2-DS-Rv	TGGTTCATCGCTTGCAGGTC
17	hph-1003-Fw	AGCCCAATAUGCTAGTGGAGGTCAACACATCA
18	hph-1003-Rv	ATTACCTAGUCGGTCGGCATCTACTCTATT
19	FW_191422	AGAGCGAUATGGAGACAGTCGCCGAGGTA
20	RV_191422	TCTGCGAUCTACAACAAAGCCCGGCGAAGGA
21	BGHA586	AGAGCGAUATGGAGGGTCCATCTCGTGTGT
22	BGHA591	TCTGCGAUTTAATTGGCCATGGCGTTGCCA
23	399441_FW	AGAGCGAUATGGTCCTTAGGACGAGGTTG
24	399441_RV	TCTGCGAUCTACAGCAATGCACGACGCA
25	191422_US_FW	GGGTTTAAUTGACTGACCCCATGGCTATGA
26	191422_US_RV	ATATTGGGCUCTTGCTAGTTGCTGAGGTACCC
27	191422_DS_FW	ACTAGGTAAUCTTGAACCTCATAGGAGTAGATAGGC
28	191422_DS_RV	GGTCTTAAUCCAATGCACCTAATCATGTCCA
29	Upst-HygR-N	CTGCTGCTCCATACAAGCCAACC
30	Dwst-1003HygF-N	GACATTGGGGAGTTCAGCGAGAG

Table A3. Cluster genes

<i>A. niger</i> enzyme	<i>A. carbonarius</i> enzyme (similarity/identity)	<i>F. graminearum</i> homologue	Predicted activity (CDD)	Predicted activity (BLAST)
50739	210466 (83.6/ 71.1)	-	Negative transcriptional regulator	hscarg dehydrogenase/NmrA family transcriptional regulator
44525	399440 (62.3/ 49.1)	AurJ	O-methyltransferase	O-methyltransferase
NapA	56260 (87.2/ 78.7)	PKS12	Polyketide synthase	Polyketide synthase
191165	56024A (88.8 / 78.4)	AurZ	EthD superfamily	Not conserved
191711	56024B (91.4 / 84.1)	-	FAD-dependent oxidoreductase	Monooxygenase
44529	399443 (86.2 / 77.4)	-	Cytochrome P450	Cytochrome P450
44530	55781 (61.6 / 54.0)	-	Not conserved	Not conserved
44531	56226 (82.3 / 68.6)	AurJ	O-methyltransferase	O-methyltransferase
44532	399445 (74.8 / 65.2)	-	Alcohol dehydrogenase, Class IV	Maleylacetate reductase
44533	8915 (62.4 / 49.2)	-	Not conserved	C6 transcription factor

Appendix E

Supporting information for Chapter 10

Molecular and chemical characterization of the biosynthesis of the 6-MSA derived meroterpenoid yanuthone D in *Aspergillus niger*

Dorte K. Holm, Lene M. Petersen, Andreas Klitgaard, Peter B. Knudsen, Zofia D. Jarczynska, Kristian F.

Nielsen, Charlotte H. Gotfredsen, Thomas O. Larsen, Uffe H. Mortensen.

SUPPORTING INFORMATION

Figures:

Figure S1. BPC of *yanA*Δ strain relative to the reference KB1001

Figure S2. The morphology of the reference strain is identical with and without addition of ¹³C₈-6-MSA

Figure S3. Positive electrospray (ESI+) mass spectra of labeled compounds

Figure S4. RT-qPCR expression analysis

Figure S5. (A) Gene deletion in *A. niger* and pDHX2

Figure S6. Base peak chromatogram (ESI+) of the five strains that express putative cluster genes *yanB*, *yanC*, *yanD*, *yanE*, and *yanG* in the *A. nidulans* IS1-*yanA* strain (Reference)

Tables:

Table S1. Detection of metabolites in the deletion and overexpression strains

Table S2. Purification of metabolites

Table S3. NMR data for yanuthone D

Table S4. NMR data for yanuthone E

Table S5. NMR data for yanuthone J

Table S6. NMR data for 7-deacetoxyyanuthone A

Table S7. NMR data for yanuthone F

Table S8. NMR data for yanuthone G

Table S9. NMR data for 22-deacetylyanuthone A

Table S10. NMR data for yanuthone H

Table S11. NMR data for yanuthone I

Table S12. NMR data for yanuthone X₁

Table S13. Fungal strains

Table S14. Primers used in the study

Table S15. ¹³C NMR chemical shift in ppm for compound 1-9 and 12

Table S16. ¹H NMR chemical shift in ppm for compound 1-9 and 12

Figure S1. BPC of *yanA* Δ strain relative to the reference KB1001, cultivated on MM (**A**), YES (**B**), CYA (**C**), and MEA (**D**) for 5 days at 30°C.

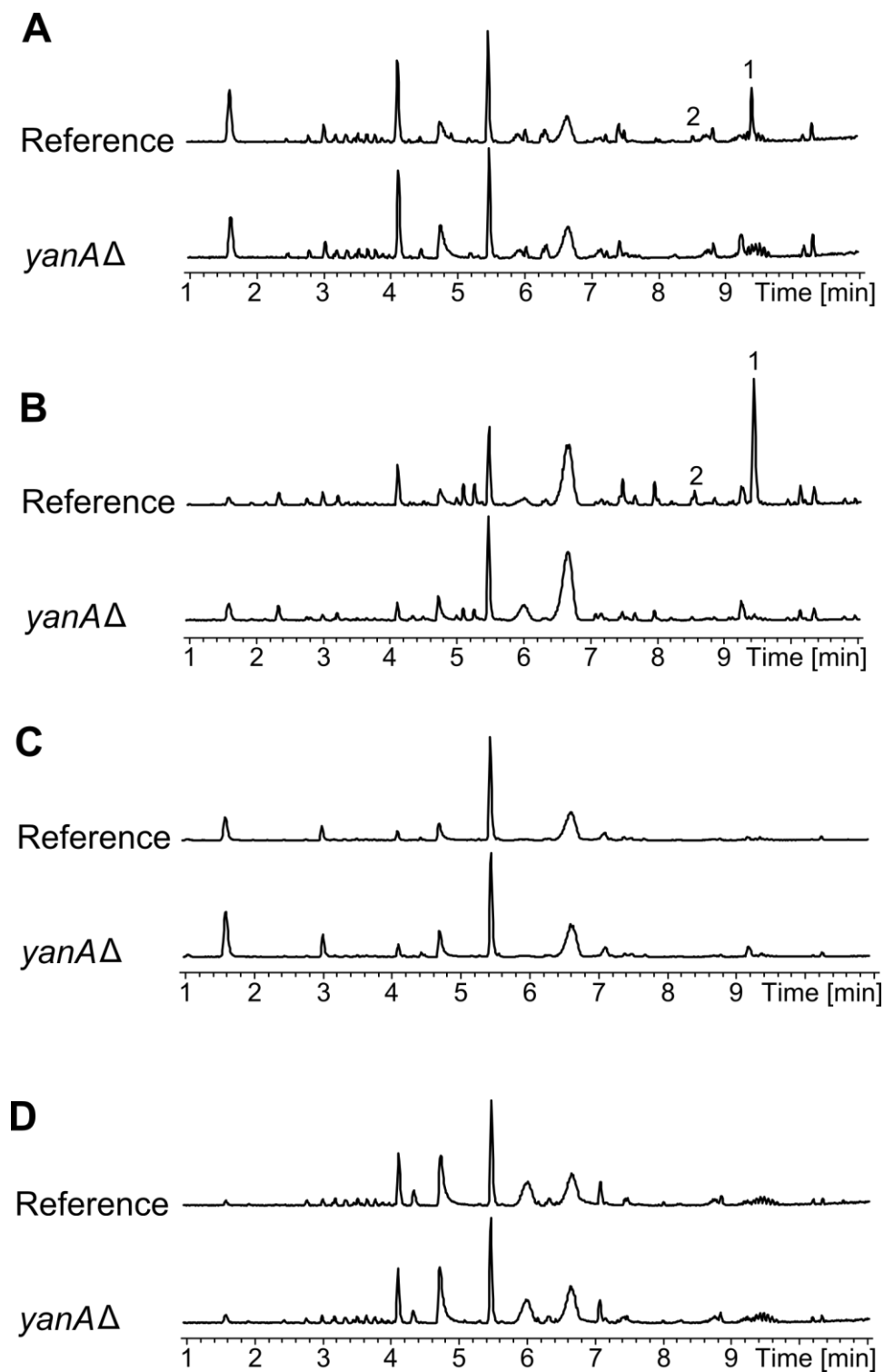


Figure S2. The morphology of the reference strain is identical with and without addition of $^{13}\text{C}_8$ -6-MSA. The figure shows the top and bottom of KB1001 cultivated on YES medium for 5 days at 30 °C.

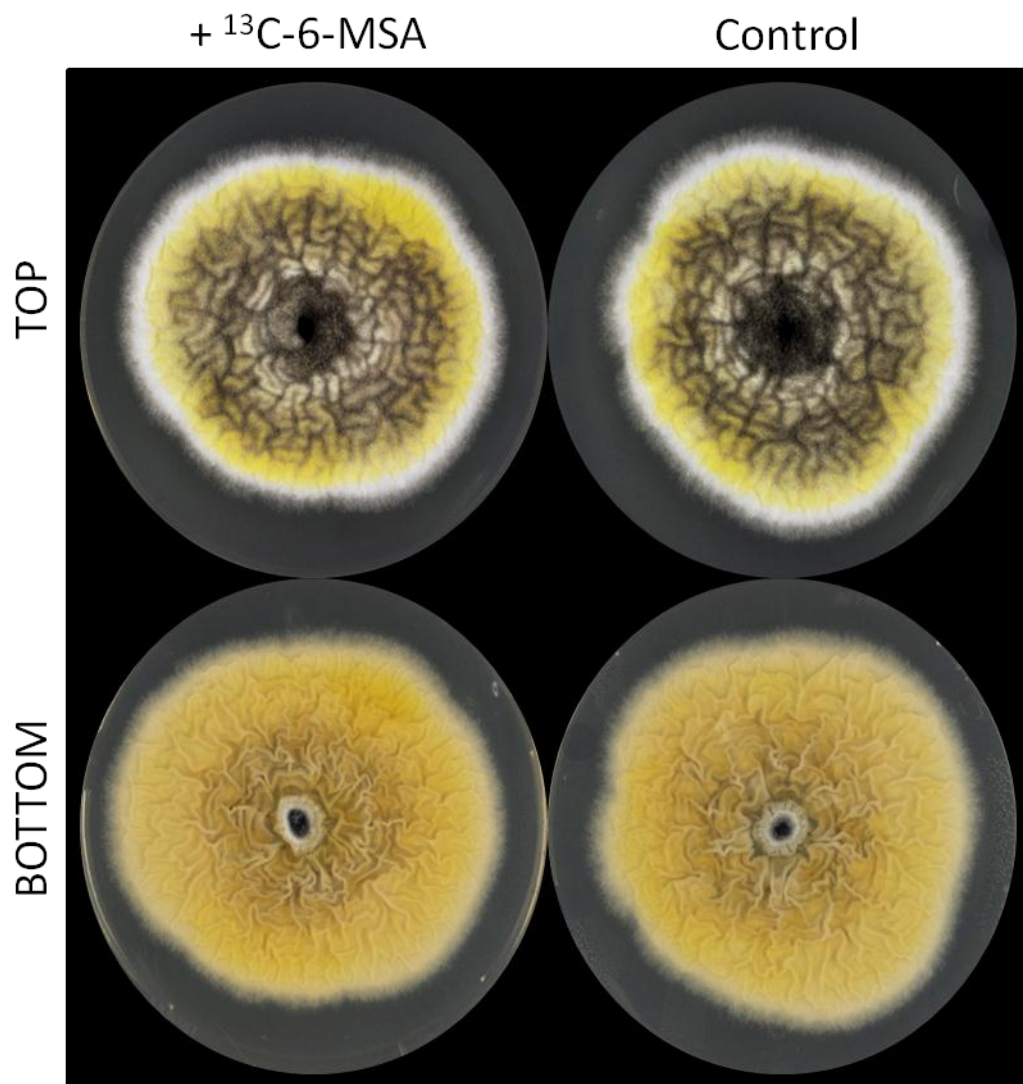


Figure S3. Positive electrospray (ESI+) mass spectrum of: **(A)** uniformly labeled $^{13}\text{C}_8$ -6-methylsalicylic acid, **(B)** unlabeled and labeled yanuthone D, **(C)** unlabeled and labeled yanuthone E, and **(D)** unlabelled yanuthone X₁. The calculated shift from $^{12}\text{C}_7$ to $^{13}\text{C}_7$ is 7.0234 Da.

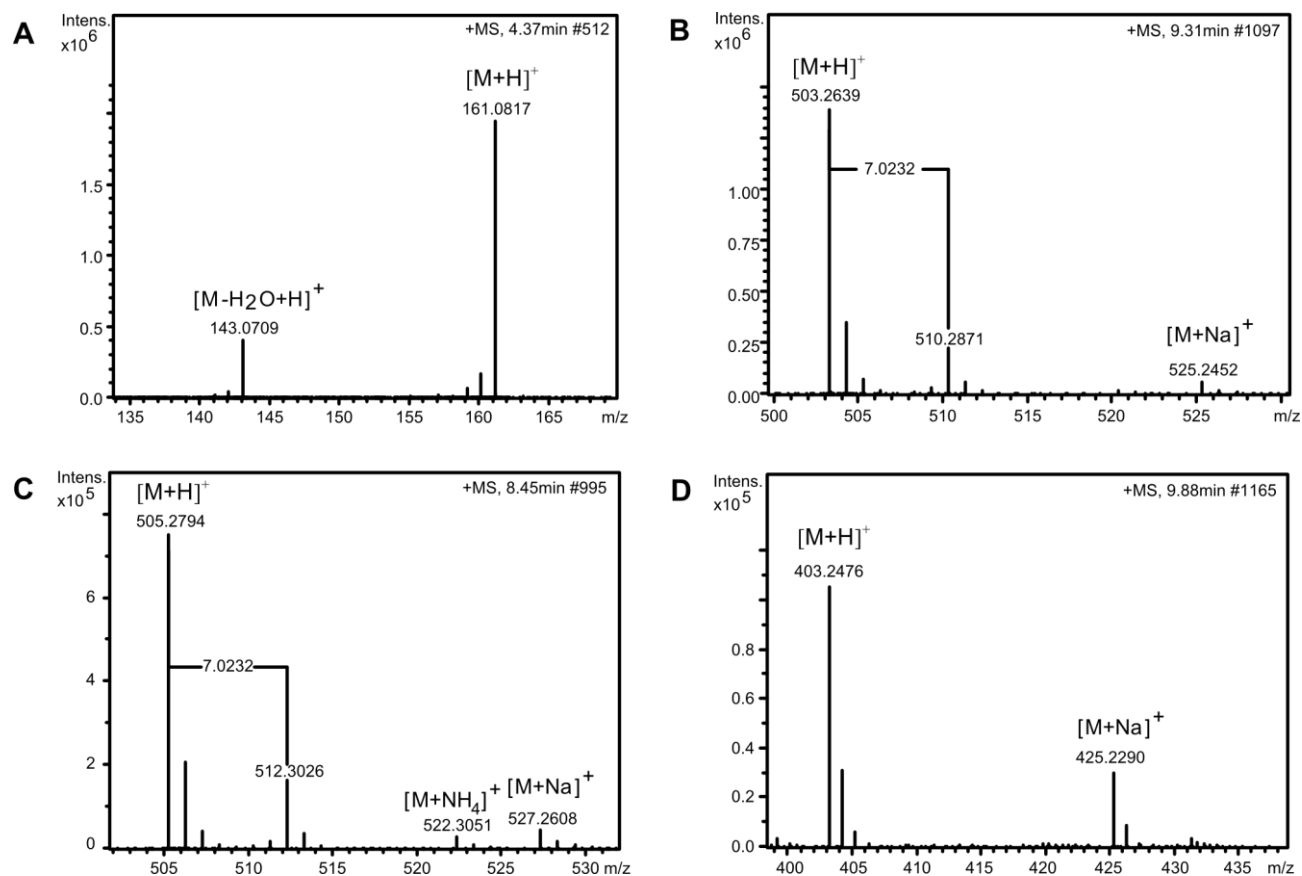


Figure S4. RT-qPCR expression analysis of *yanA* and 13 flanking genes in a *yanR* Δ (44961 Δ) strain and a reference strain. **A.** Absolute expression levels ($(\Delta c(t))^{-1}$ values) in the reference strain (grey bar) compared to the TF Δ strain (blue/red bar). Red bars indicate significant down-regulation (p-value <0.05), blue bars indicate non-significant changes (p-value >0.05). Values are normalized to expression of *hhtA*. **B.** Relative expression levels (fold change, $2^{\Delta\Delta c(t)}$ values) of *yan* cluster genes (except *yanR*, which is deleted) in the TF Δ strain relative to the reference strain. Expression levels are individually normalized to expression of *actA* and *hhtA*, as indicated.

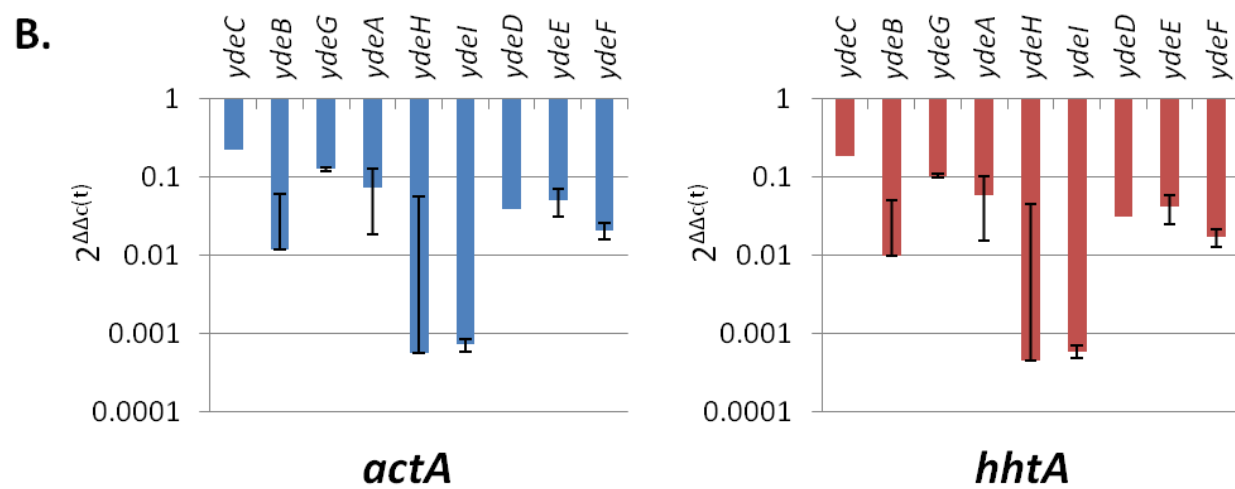
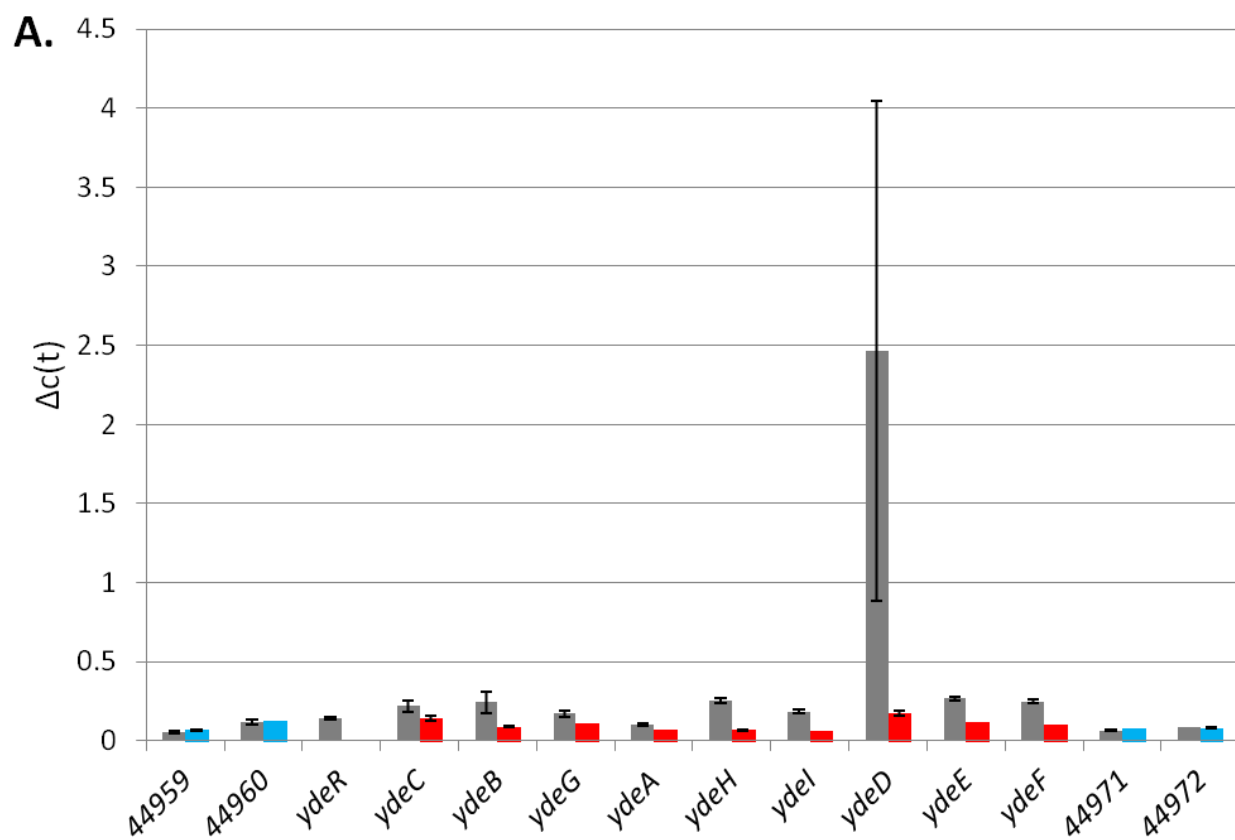


Figure S5. (A) Gene deletion in *A. niger*. The ORF is replaced by the selectable hygromycin B phosphatase (*hph*) gene. Not to scale. **(B)** pDHX2 vector used for expression of all cluster genes in the *A. nidulans* IS1-*yanA* strain. The vector contains the AMA1 sequence for autonomous replication in *Aspergillus* and the *pyrG* gene for selection.

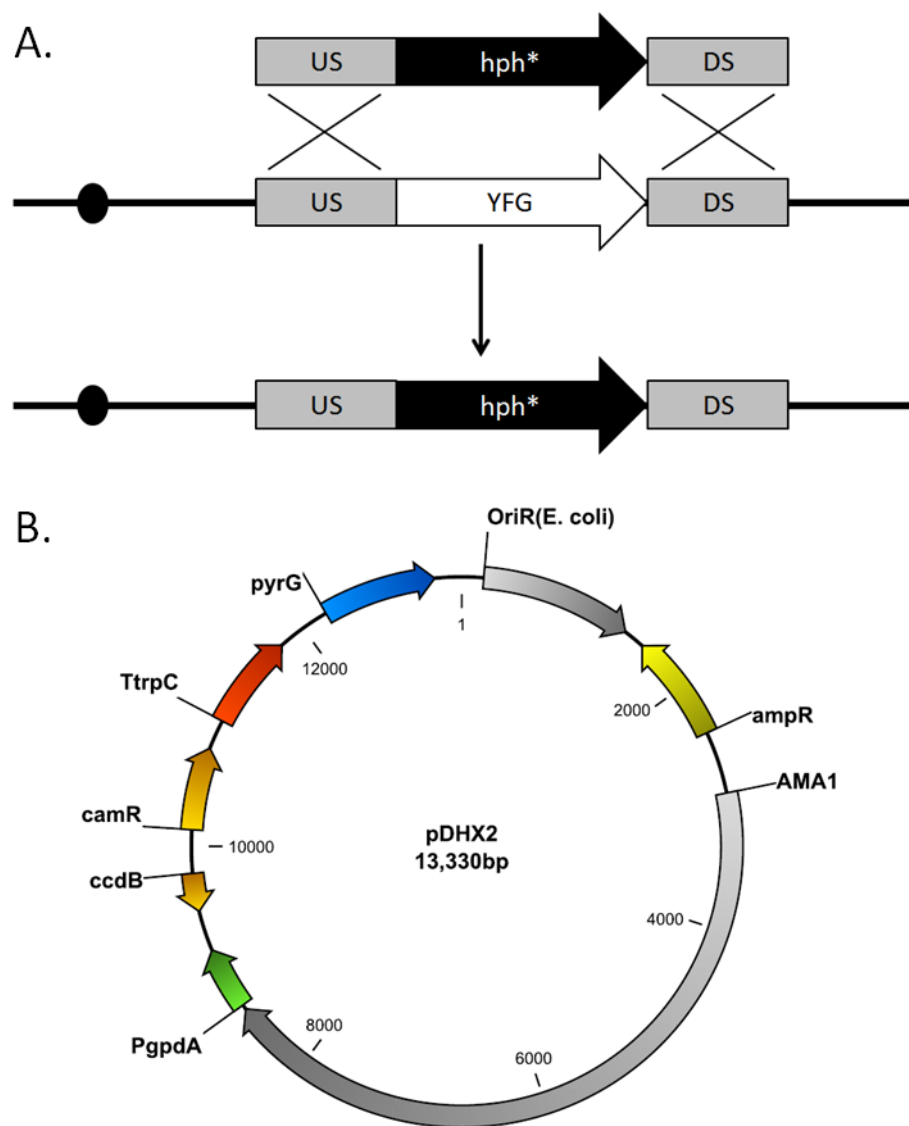


Figure S6. Base peak chromatogram (ESI+) of the five strains that express putative cluster genes *yanB*, *yanC*, *yanD*, *yanE*, and *yanG* in the *A. nidulans* IS1-*yanA* strain (Reference). All strains, except Oex-*yanB* still produce 6-MSA. IS = internal standard: chloroamphenicol, A = austinol, DHA = dehydroaustinol. Chromatograms are to scale.

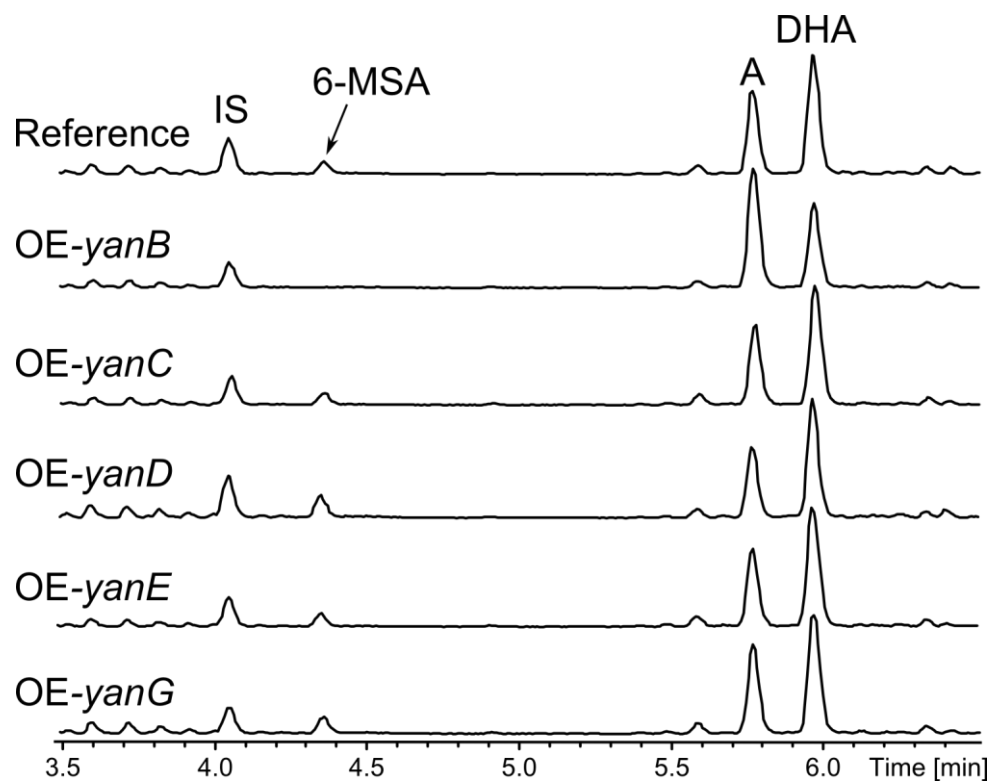


Table S1. Detection of metabolites in the deletion and overexpression strains using extracted ion chromatograms of $[M+H]^+ \pm 0.005$.

	6-MSA	Yanuthone D (1)	Yanuthone E (2)	7-deacetoxyyanuthone A (3)	Yanuthone F (4)	Yanuthone G (5)	22-deacetyl Yanuthone A (6)	Yanuthone H (7)	Yanuthone I (8)	Yanuthone J (9)	Yanuthone X ₁ (12)
WT	-	+	+	-	-	+	-	-	+	-	+
OE- <i>yanA</i>	+	-	-	-	-	-	-	-	-	-	-
<i>yanA</i> Δ	-	-	-	-	-	-	-	-	-	-	+
<i>yanB</i> Δ	-	-	-	-	-	+	-	-	+	-	+
<i>yanC</i> Δ	-	-	-	-	-	-	-	-	-	-	-
<i>yanD</i> Δ	-	-	-	-	-	-	-	-	-	-	-
<i>yanE</i> Δ	-	-	-	-	-	-	-	-	-	-	-
<i>yanF</i> Δ	-	-	+	+	-	-	-	-	+	+	+
<i>yanG</i> Δ	-	-	-	-	-	-	-	-	-	-	-
<i>yanH</i> Δ	-	-	-	+	+	+	-	-	-	-	+
<i>yanI</i> Δ	-	-	-	+	-	-	+	+	+	-	+
<i>yanR</i> Δ	-	-	-	+	-	+	-	-	-	-	+

Structural elucidation.

The structural elucidation of the compounds showed several similar features in the ^1H as well as 2D spectra comparable to those reported for the known yanuthones.[Bugni *et al.*, 2000; Li *et al.*, 2003] All compounds except yanuthone I displayed 8H overlapping resonances at δ_{H} 1.93-2.11 ppm in the ^1H spectrum corresponding to the four methylene groups H4, H5, H8 and H9 in the sesquiterpene moiety. Other common resonances were from the diastereotopic pair H-12/H-12' and 3 methyl groups (H-19, H-20 and H-21) around δ_{H} 1.60 ppm, whereof H-20 and H-21 were overlapping. In the HMBC spectrum a correlation to the quaternary C-18 around δ_{C} 194 ppm was seen and all compounds also had two carbons around 60 ppm (one quaternary, one methine) being the carbons in the epoxide ring.

The compounds however differed greatly in the moiety attached to C-16. Yanuthone D, yanuthone E and yanuthone J all displayed two methylene groups around δ_{C} 45 ppm, a methyl group around δ_{C} 28 ppm, two carbonyls around δ_{C} 171 ppm and another quaternary carbon around δ_{C} 70 ppm for the mevalonic acid part. 7-deacetoxyyanuthone A, yanuthone F and yanuthone G all had a methyl group attached at C-16 while yanuthone H, yanuthone I and 22-deacetylyanuthone A had a further hydroxy group at C-22. The hydroxylation in this position was indicated by a significant shift downfield of H-22/C-22.

Some structures had a further modification being a hydroxy group at either C-1 or C-2. The compounds yanuthone F, G and H all had a hydroxy group at C-1, shifting C-1 and H-1 significantly downfield. Yanuthone J had the hydroxy group attached at C-2 which shifted the resonances for C-2 and H-2 downfield, and due to the lack of the double bond in those structures, the resonance for H-3 was no longer observed in the double bond area but at δ_{H} 1.35 ppm. Yanuthone I differed in this part of the structure with fewer resonances due to the shorter terpene chain.

The ^1H NMR spectrum for yanuthone G stood out from the rest due to several resonances between 3-5 ppm. Elucidation of the structure revealed a sugar moiety attached to the hydroxy group at C-15. The presence of this hexose unit gave rise to the additional resonances observed.

The NMR data for yanuthone X₁ displayed the same resonances for the sesquiterpene part of the molecule, but the methoxy group attached to C-16 is different for all other reported yanuthones, and was obvious from the chemical shift of C-16 which gave rise to a resonance at δ_{C} 168.3 ppm, which is considerable further downfield than in the other structures. Furthermore C-17 was affected shifting upfield to δ_{C} 100.3 ppm.

NMR data for all compounds can be found in Tables S3-S12 and a combined overview of the chemical shifts in Tables S15 and S16.

The stereochemistry of the compounds was investigated by circular dichroism (CD) and optical rotation. The CD data for yanuthone D, E and 7-deacetoxyyanuthone A showed that the positive and negative cotton effects were identical to those previously reported for these compounds [Bugni *et al.*, 2000].

Table S2. Purification of metabolites

Compound name	Strain	Number of plates	Extraction	Isolera fractionation	Purification	Yield
Yanuthone D (1)	<i>A. niger</i> KB1001	200	EtOAc + 1 % FA	50g diol, 40 mL·min ⁻¹ , CV = 66 mL, auto fractionation, 15 fractions	Waters, Luna II C ₁₈ , 4 mL·min ⁻¹ , 40-100% ACN over 20 min.	4.5 mg
Yanuthone E (2)	<i>A. niger</i> KB1001	200	EtOAc + 1 % FA	50g diol, 40 mL·min ⁻¹ , CV = 66 mL, auto fractionation, 15 fractions	Gilson, Luna II C ₁₈ , 5 mL·min ⁻¹ , 40-100% ACN over 20 min.	2.9 mg
7-deacetoxy Yanuthone A (3)	<i>A. niger</i> KB1001	200	EtOAc + 1 % FA	50g diol, 40 mL·min ⁻¹ , CV = 66 mL, auto fractionation, 15 fractions	Waters, Luna II C ₁₈ , 4 mL·min ⁻¹ , 40-100% ACN over 20 min.	9.3 mg
Yanuthone F (4)	<i>yanHΔ</i>	100	EtOAc	25g diol, 25 mL·min ⁻¹ , CV = 33 mL, auto fractionation, 10 fractions	Gilson, Luna II C ₁₈ , 5 mL·min ⁻¹ , 30-80% ACN over 18 min. Waters, gemini, 5 mL·min ⁻¹ , 40-65% ACN over 20 min	1.8 mg
Yanuthone G (5)	<i>yanHΔ</i>	100	EtOAc	25g diol, 25 mL·min ⁻¹ , CV = 33 mL, auto fractionation, 10 fractions	Gilson, Luna II C ₁₈ , 5 mL·min ⁻¹ , 30-80% ACN over 18 min. Waters, gemini, 5 mL·min ⁻¹ , 30-40% ACN over 20 min, to 45% for 2 min.	4.0 mg
22-deacetylyanuthone A (6)	<i>yanIΔ</i>	100	EtOAc	25g diol, 25 mL·min ⁻¹ , CV = 33 mL, auto fractionation, 12 fractions	Gilson, Luna II C ₁₈ , 5 mL·min ⁻¹ , 20-90% ACN over 17 min. Waters, Luna II C ₁₈ , 5 mL·min ⁻¹ , 30-100% ACN over 20 min.	7.3 mg
Yanuthone H (7)	<i>yanIΔ</i>	100	EtOAc	25g diol, 25 mL·min ⁻¹ , CV = 33 mL, auto fractionation, 12 fractions	Gilson, Luna II C ₁₈ , 5 mL·min ⁻¹ , 30-60% ACN over 16 min.	9.8 mg
Yanuthone I (8)	<i>yanIΔ</i>	100	EtOAc	25g diol, 25 mL·min ⁻¹ , CV = 33 mL, auto fractionation, 12 fractions	Waters, Luna II C ₁₈ , 5 mL·min ⁻¹ , 30-60% ACN over 16 min.	4.7 mg
Yanuthone J (9)	<i>yanFΔ</i>	150	EtOAc + 1 % FA	25g diol, 25 mL·min ⁻¹ , CV = 33 mL, auto fractionation, 15 fractions	Waters, Luna II C ₁₈ , 4 mL·min ⁻¹ , 20-100% ACN over 20 min.	1.5 mg
Yanuthone X ₁ (12)	<i>A. niger</i> KB1001	200	EtOAc + 1 % FA	50g diol, 40 mL·min ⁻¹ , CV = 66 mL, auto fractionation, 15 fractions	Waters, gemini, 4 mL·min ⁻¹ , 90 % ACN isocratic for 15 min, then to 100 % ACN for 5 min.	1.5 mg

Yanuthone D

HRESIMS: $m/z = 503.2640$ $[M + H]^+$, calculated for $[C_{28}H_{38}O_8 + H]^+$: 503.2639. $[\alpha]_{587}^{20} = 32.4^\circ$

CD_{MeOH}: $[\theta]_{213} = 12.4263$, $[\theta]_{230} = -7.8722$, $[\theta]_{240} = 1.1769$, $[\theta]_{260} = -4.0282$, $[\theta]_{298} = 0.7690$

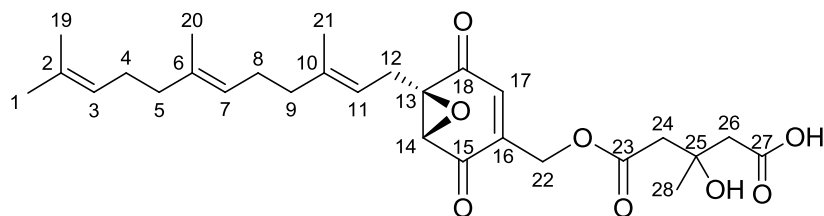


Table S3. NMR data for yanuthone D

Atom assignment	¹ H-chemical shift [ppm]/ J coupling constants [Hz]	¹³ C-chemical shift [ppm]	HMBC correlations	NOESY connectivities
1	1.59 (3H, s)	17.1	2, 19	-
2	-	132.1	-	-
3	5.10 (1H, m)	125.0	1, 19	19
4	2.06 (2H, m)	27.3	2, 3, 5/9	20
5	1.97 (2H, m)	40.2	3/7, 6, 4/8,20	20
6	-	136.4	-	-
7	5.09 (1H, m)	124.9	5/9, 8, 20	9
8	2.09 (2H, m)	27.0	5/9, 6, 7, 10	20, 21
9	1.99 (2H, m)	40.1	7, 8, 10, 11, 21	7, 11
10	-	141.0	-	-
11	5.03 (1H, t, 1.5)	116.4	9, 12, 21	9, 12, 12', 14
12	2.66 (1H, m)	25.6	10, 11, 13, 14, 15/18	11, 12', 14, 21
12'	2.76 (1H, m)	25.6	10, 11, 13, 14, 15/18	11, 12, 14, 21
13	-	63.0	-	-
14	3.68 (1H, s)	58.6	12, 13, 15/18, 16, 22	11, 12, 12', 21
15	-	193.2	-	-
16	-	143.9	-	-
17	6.58 (1H, t, 1.5)	133.1	13, 15/18, 22	22, 22', 24, 28
18	-	193.2	-	-
19	1.66 (3H, s)	25.7	1, 2, 3	3, 4
20	1.591 (3H, s)	16.1	5, 6, 7	5, 8
21	1.63 (3H, s)	16.2	9, 10, 11	8, 12, 12', 14
22	4.84 (1H, dd, 16.1, 1.5)	60.2	15/18, 16, 17, 23	17
22'	4.89 (1H, dd, 16.1, 1.5)	60.2	15/18, 16, 17, 23	17
23	-	171.2	-	-
24	2.69 (2H, m)	45.6	23, 26	17, 28
25	-	70.0	-	-
26	2.64 (1H, m)	45.1	24, 25, 28	28
26'	2.57 (1H, m)	45.1	27	-
27	-	173.1	-	-
28	1.33 (3H, s)	27.6	24, 25	17, 24, 26

Yanuthone E

HRESIMS: $m/z = 505.2791$ $[M + H]^+$, calculated for $[C_{28}H_{40}O_8 + H]^+$: 505.2789. $[\alpha]_{587}^{20} = 11.7^\circ$

CD_{MeOH}: $[\theta]_{209} = 2.1815$, $[\theta]_{241} = -12.3941$, $[\theta]_{340} = 7.4622$

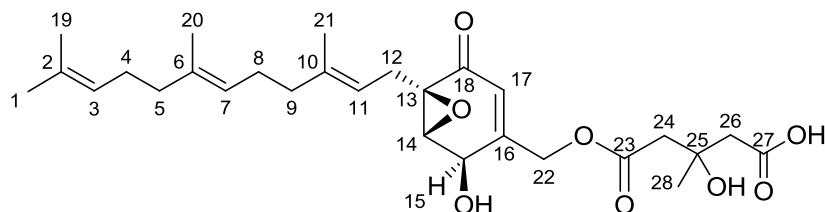


Table S4. NMR data for yanuthone E

Atom assignment	¹ H-chemical shift [ppm]/ J coupling constants [Hz]	¹³ C-chemical shift [ppm]	HMBC correlations	NOESY connectivities
1	1.58 (3H, s)	17.7	2, 3, 19	-
2	-	131.8	-	-
3	5.07 (1H, m)	124.9	1, 4, 5, 19	4, 5, 19
4	2.02 (2H, m)	27.2	2, 6	3
5	1.96 (2H, m)	40.3	3/7, 4, 6, 20,	3, 7
6	-	135.8	-	-
7	5.06 (1H, t, 6.4)	124.9	5/9, 8, 20	5, 9, 20
8	2.06 (2H, m)	27.2	5/9, 6, 7, 10	-
9	1.99 (2H, m)	40.3	6, 7, 8, 11, 21	7, 11
10	-	139.9	-	-
11	5.04 (1H, t, 7.2)	117.7	9, 12, 13, 21	9, 12, 12', 14
12	2.67 (1H, m)	26.6	10, 11, 13, 14, 18	11, 12', 14
12'	2.42 (1H, m)	26.6	10, 11, 13, 14, 18	11, 12, 14
13	-	61.3	-	-
14	3.64 (1H, d, 2.8)	60.1	12/12', 13, 15, 16, 22/22',	11, 12, 12' 15
15	4.68 (1H, br. s)	65.8	16, 17	14
16	-	154.6	-	-
17	5.86 (1H, q, 1.4)	121.6	13, 15, 16, 22, 22'	22, 22', 24
18	-	194.8	-	-
19	1.66 (3H, s)	25.6	1, 2, 3	3
20	1.581 (3H, s)	16.0	5, 6, 7	7
21	1.62 (3H, s)	16.3	9, 10, 11	-
22	4.82 (1H, d, 16.1)	63.6	15, 16, 17, 18, 23	17, 22'
22'	4.76 (1H, d, 16.1)	63.6	15, 16, 17, 18, 23	17, 22
23	-	171.4	-	-
24	2.69 (2H, m)	45.8	23, 25, 26, 28	17
25	-	70.2	-	-
26	2.59 (2H, m)	45.4	24, 25, 27, 28	-
27	-	173.5	-	-
28	1.33 (3H, s)	27.7	25, 26	-

Yanuthone J

HRESIMS: $m/z = 523.2907$ $[M + H]^+$, calculated for $[C_{28}H_{42}O_9 + H]^+$: 523.2901 $[\alpha]_{587}^{20} = 3.6^\circ$

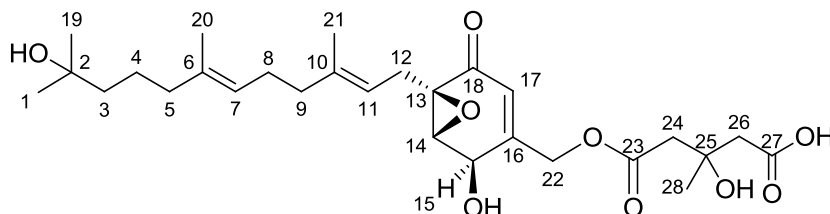


Table S5. NMR data for yanuthone J

Atom assignment	^1H -chemical shift [ppm]/ J coupling constants [Hz]	^{13}C -chemical shift [ppm]	HMBC correlations
1	1.28 (3H, m)	29.7	-
2	-	70.9	-
3	1.35 (2H, m)	44.0	4, 19
4	1.43 (2H, m)	23.1	-
5	1.93 (2H, m)	40.5	3, 4, 6, 7, 11, 20
6	-	136.3	-
7	5.08 (1H, tq, 6.3, 1.0)	124.4	-
8	2.09 (2H, m)	26.7	6, 7, 9, 10
9	2.02 (2H, m)	40.3	7, 8, 10, 11, 21
10	-	139.6	-
11	5.03 (1H, tq, 6.3, 1.0)	117.5	-
12	2.45 (1H, m)	26.5	10, 11, 13, (18)
12'	2.68 (1H, m)	26.5	10, 11, 13
13	-	61.1	-
14	3.63 (1H, d, 2.8)	59.9	13, 15, 16
15	4.69 (1H, m)	65.7	-
16	-	154.7	-
17	5.86 (1H, q, 1.5)	121.5	13, 15, 22
18	-	(194.2)	-
19	1.12 (3H, s)	29.2	1, 2, 3
20	1.58 (3H, s)	15.9	5, 6, 7
21	1.62 (3H, s)	16.2	9, 10, 11
22	4.83 (1H, d, 16.0)	63.6	16, 17, 23
22'	4.76 (1H, d, 16.0)	63.6	16, 17, 23
23	-	171.2	-
24	2.681 (2H, m)	43.7	23, 25, 26, 28
25	-	70.9	-
26	2.61 (2H, m)	45.1	24, 25, 27, 28
27	-	173.6	-
28	1.33 (3H, s)	27.5	25, 26

7-deacetoxyyanuthone A

HRESIMS: $m/z = 345.2434$ $[M + H]^+$, calculated for $[C_{22}H_{32}O_3 + H]^+$: 345.2424 $[\alpha]_{587}^{20} = 20.0^\circ$

CD_{MeOH}: $[\theta]_{206} = -13.6324$, $[\theta]_{242} = -25.0218$, $[\theta]_{335} = 16.7472$

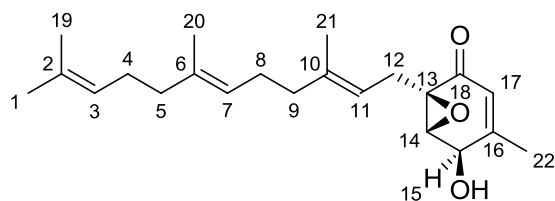


Table S6. NMR data for 7-deacetoxyyanuthone A

Atom assignment	¹ H-chemical shift [ppm]/ J coupling constants [Hz]
1	1.67 (3H, d 1.0)
2	-
3	5.10 (1H, m)
4	2.07-1.90 (2H, m)
5	2.07-1.90 (2H, m)
6	-
7	5.09 (1H, m)
8	2.07-1.90 (2H, m)
9	2.07-1.90 (2H, m)
10	-
11	5.05 (1H, m)
12	2.72 (1H, dd, 15.2, 7.9)
12'	2.39 (1H, dd, 15.2, 6.7)
13	-
14	3.62 (1H, d, 2.89)
15	4.49 (1H, br. s)
16	-
17	5.70 (1H, p, 1.5)
18	-
19	1.60 (3H, s)
20	1.59 (3H, s)
21	1.64 (3H, s)
22	2.07-1.90 (3H, m)

Yanuthone F

HRESIMS: $m/z = 361.2373$ $[M + H]^+$, calculated for $[C_{22}H_{32}O_4 + H]^+$: 361.2373 $[\alpha]_{587}^{20} = 14.2^\circ$

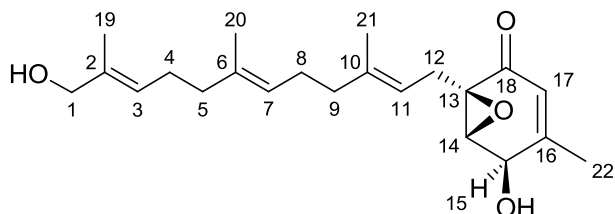


Table S7. NMR data for yanuthone F

Atom assignment	^1H -chemical shift [ppm]/ J coupling constants [Hz]	^{13}C -chemical shift [ppm]	HMBC correlations	NOESY connectivities
1	3.84 (2H, br. s)	68.3	2, 3, 19	3, 5
2	-	136.2	-	-
3	5.33 (1H, tq, 7.4, 1.5)	125.3	1, 19	1, 4, 5
4	2.11 (2H, m)	26.8	2, 5, 6	3, 7
5	1.99 (2H, m)	40.1	2/6, 4/8, 20	1, 3, 7
6	-	136.2	-	-
7	5.09 (1H, tq, 6.9, 1.1)	124.9	4/8, 5/9, 20	4, 5, 8, 9
8	2.07 (2H, m)	26.8	6, 5/9, 10	7
9	2.01 (2H, m)	40.1	6, 8/12, 10, 21	7, 11, 14
10	-	139.7	-	-
11	5.04 (1H, tq, 7.3, 1.2)	118.0	8/12, 9, 21	9, 12, 12', 14
12	2.70 (1H, dd, 15.1, 8.1)	27.0	10, 11, 13	11, 12', 14
12'	2.37 (1H, m)	27.0	10, 11, 13, 18	11, 12, 14
13	-	61.4	-	-
14	3.60 (1H, d, 2.7)	60.2	13, 15, 16	9, 11, 12, 12', 15
15	4.48 (1H, m)	67.6	16, 17	14, 22
16	-	158.8	-	-
17	5.68 (1H, m)	123.2	13, 15, 22	22
18	-	194.9	-	-
19	1.59 (3H, s)	13.5	1, 2, 3, 4	-
20	1.591 (3H, s)	16.1	5, 6, 7	-
21	1.62 (3H, s)	16.1	6, 10, 11, 13	-
22	1.92 (2H, br. s)	20.0	15, 16	15, 17

Yanuthone G

HRESIMS: $m/z = 523.2917$ $[M + H]^+$, calculated for $[C_{28}H_{42}O_9 + H]^+$: 523.2901. $[\alpha]_{587}^{20} = 19.6^\circ$

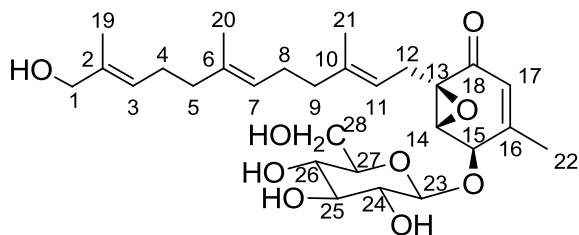


Table S8. NMR data for yanuthone G

Atom assignment	^1H -chemical shift [ppm]/ J coupling constants [Hz]	^{13}C -chemical shift [ppm]	HMBC correlations	NOESY connectivities
1	3.86 (2H, s)	68.2	2, 3, 19	3, 19
2	-	135.9	-	-
3	5.34 (1H, t, 6.3)	125.2	1, 4, 5, 19	1
4	2.10 (2H, m)	26.8	2/6, 3, 5	5, 20
5	1.99 (2H, m)	40.0	2/6, 4/8, 7, 20	4, 7
6	-	135.9	-	-
7	5.10 (1H, t, 6.3)	124.9	4/8, 5/9, 20	5, 9
8	2.06 (2H, m)	26.9	5/9, 6, 10	9, 20, 21
9	2.02 (2H, m)	40.0	6, 8/12, 10, 11, 21	7, 8, 11, 14
10	-	139.7	-	-
11	5.04 (1H, t, 6.7)	117.9	8/12, 9, 21	9, 12, 12', 14, 15
12	2.67 (1H, m)	26.9	10, 11, 13, 18	11, 14, 21
12'	2.37 (1H, dd, 14.8, 6.2)	26.9	10, 11, 13, 18	11, 14, 21
13	-	61.3	-	-
14	3.82 (1H, br. s)	59.7	15, 16	9, 11, 12, 12', 15, 17, 21
15	4.59 (1H, m)	76.2	23	11, 14, 17
16	-	157.0	-	-
17	5.71 (1H, br. s)	123.9	13, 15, 22	14, 15, 27
18	-	194.4	-	-
19	1.591 (3H, s)	13.7	1, 2, 3	1
20	1.59 (3H, s)	15.9	5, 6, 7	4, 8
21	1.62 (3H, s)	16.3	9, 10, 11	8, 12, 12', 14
22	1.97 (3H, m)	20.1	15, 16	26, 27
23	4.56 (1H, m)	105.7	16	24, 25, 26, 27, 28, 28'
24	3.25 (1H, br. s)	74.6	23, 26/27	23, 25, 26, 27
25	3.37 (1H, m)	77.2	-	23, 24, 26, 27, 28, 28'
26	3.35 (1H, m)	71.2	-	22, 23, 24, 25, 27, 28, 28'
27	3.35 (1H, m)	77.2	-	17, 22, 23, 24, 25, 26, 28, 28'
28	6.77 (1H, d, 10.7)	62.6	-	23, 25, 26, 27, 28'
28'	3.66 (1H, d, 10.7)	62.6	-	23, 25, 26, 27, 28

22-deacetylanuthone A

HRESIMS: $m/z = 361.2372$ $[M + H]^+$, calculated for $[C_{22}H_{32}O_4 + H]^+$: 361.2373 $[\alpha]_{587}^{20} = 30.6^\circ$

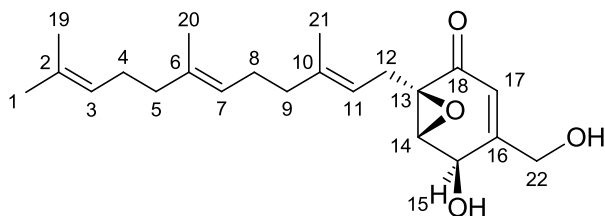


Table S9. NMR data for 22-deacetylanuthone A

Atom assignment	^1H -chemical shift [ppm]/ J coupling constants [Hz]	^{13}C -chemical shift [ppm]	HMBC correlations	NOESY connectivities
1	1.66 (3H, s)	25.8	2, 3, 19	-
2	-	132.0	-	-
3	5.10 (1H, m)	124.9	1, 4, 5, 19	4, 5, 19
4	2.07 (2H, m)	26.8	2, 5, 6	3, 5, 19, 20
5	1.98 (2H, m)	40.1	4/8, 6, 3/7, 20	3, 4, 7
6	-	136.0	-	-
7	5.10 (1H, m)	124.9	4/8, 5/9, 20	8, 5/9, 20
8	2.09 (2H, m)	26.8	5/9, 6, 7, 10	7, 9, 20, 21
9	2.03 (2H, m)	40.1	7, 10, 11, 13, 8/12, 21	7, 8, 11
10	-	140.2	-	-
11	5.05 (1H, ddd, 7.93, 6.71, 1.22)	118.0	9, 8/12, 13, 21	9, 12, 12', 14, 21
12	2.70 (1H, dd, 15.3, 7.9)	26.8	10, 13, 17, 18	11, 12', 14, 21
12'	2.42 (1H, dd, 15.3, 6.7)	26.8	10, 13, 17, 18	11, 12, 14, 21
13	-	61.6	-	-
14	3.62 (1H, d, 2.8)	60.2	12, 15, 16, 22	11, 12, 12', 15
15	4.64 (1H, m)	66.2	13, 16, 17, 18	14, 22, 22'
16	-	160.8	-	-
17	5.87 (1H, q, 1.73)	119.7	15, 16, 22	22, 22'
18	-	195.1	-	-
19	1.601 (3H, s)	17.7	1, 2, 3	3, 4
20	1.60 (3H, s)	15.9	5, 6, 7	4, 7, 8
21	1.63 (3H, s)	16.2	8/12, 9, 10, 11, 13	8, 11, 12, 12'
22	4.27 (1H, m)	62.0	15	15, 17
22'	4.21 (1H, m)	62.0	-	15, 17

Yanuthone H

HRESIMS: $m/z = 377.2332$ $[M + H]^+$, calculated for $[C_{22}H_{32}O_5 + H]^+$: 377.2322. $[\alpha]_{587}^{20} = 27.7^\circ$

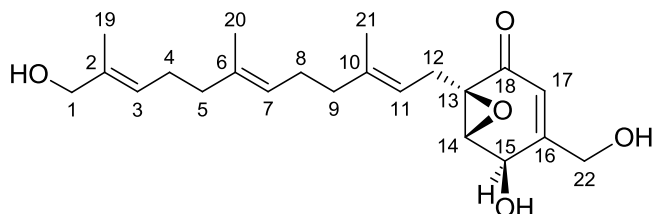


Table S10. NMR data for yanuthone H

Atom assignment	^1H -chemical shift [ppm]/ J coupling constants [Hz]	^{13}C -chemical shift [ppm]	HMBC correlations	NOESY connectivities
1	3.85 (2H, s)	68.3	2, 3, 4, 19	3
2	-	136.0	-	-
3	5.34 (1H, m)	125.4	1, 4, 5, 19	1, 5
4	2.11 (2H, m)	26.8	3, 5, 6	-
5	1.99 (2H, m)	40.1	4/8, 6, 7, 12, 20	3
6	-	135.9	-	-
7	5.10 (1H, t, 6.4)	124.9	4/8, 5/9, 20	9
8	2.07 (2H, m)	26.8	5/9, 7, 10	-
9	2.02 (2H, m)	40.1	7, 10, 11, 8/12, 21	7, 11
10	-	139.8	-	-
11	5.05 (1H, t, 6.8)	118.0	8/12, 21	9, 12, 12', 14
12	2.71 (1H, dd, 15.1, 8.3)	26.8	10, 11, 13, 18	11, 12', 14
12'	2.40 (1H, dd, 15.1, 6.8)	26.8	20, 11, 13, 18	11, 12, 14
13	-	60.9	-	-
14	3.62 (1H, d, 2.9)	60.4	12, 13, 15, 16	11, 12, 12' 15
15	4.64 (1H, br. s)	66.1	16, 17	14, 16, 17
16	-	160.6	-	15
17	5.87 (1H, d, 1.5)	119.7	13, 15, 16	15, 22, 22'
18	-	194.9	-	-
19	1.601 (3H, s)	13.5	1, 2	-
20	1.60 (3H, s)	16.1	5/8, 7	-
21	1.63 (3H, s)	16.2	8/12, 10, 11	-
22	4.27 (1H, d, 17.6)	62.0	15, 16, 17, 18	17, 22'
22'	4.21 (1H, m)	62.0	15, 16, 17, 18	17, 22

Yanuthone I

HRESIMS: $m/z = 325.1635$ $[M + H]^+$, calculated for $[C_{17}H_{24}O_6 + H]^+$: 325.1646. $[\alpha]_{587}^{20} = 21.1^\circ$

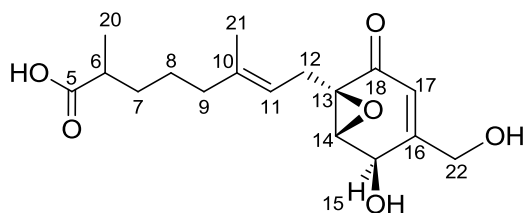


Table S11. NMR data for yanuthone I

Atom assignment	^1H -chemical shift [ppm]/ J coupling constants [Hz]	^{13}C -chemical shift [ppm]	HMBC correlations	NOESY connectivities
5	-	177.4	-	-
6	2.28 (1H, m)	38.2	5, 7/7', 8, 20	7, 7', 8, 20
7	1.48 (1H, m)	32.5	5, 6, 8, 20	6, 7', 8, 9, 20
7'	1.26 (1H, m)	32.5	5, 6, 8, 20	6, 7, 8, 9, 21
8	1.32 (2H, m)	24.5	7/7', 9, 10	6, 7, 7', 9, 11, 20
9	1.92 (2H, t, 7.1)	38.7	7/7', 8, 10, 11, 21	7, 7', 8, 11, 14
10	-	137.9	-	-
11	5.00 (1H, t, 7.1)	116.9	9, 12/12', 13, 21	8, 9, 12, 12', 14
12	2.63 (1H, dd, 15.1, 7.9)	25.7	10, 11, 13, 14, 17, 18	12', 14, 17, 21
12'	2.32 (1H, m)	25.7	9, 10, 11, 13, 14, 17, 18	12, 14, 17, 21
13	-	60.0	-	-
14	3.59 (1H, d, 2.4)	59.1	12/12', 13/22/22', 15, 16	9, 11, 12, 12', 15, 17, 21
15	4.61 (1H, br. S)	64.3	16, 17	14, 22'
16	-	162.2	-	-
17	5.82 (1H, d, 1.6)	117.4	13/22/22', 14, 15, 16, 18	22'
18	-	193.7	-	-
20	1.02 (3H, d, 6.7)	16.7	5, 6, 7	6, 7, 7', 8
21	1.56 (3H, s)	15.6	9, 10, 11, 12/12', 13	7', 9, 12, 12', 14
22	4.21 (1H, d, 18.5)	60.0	15, 16, 17, 18	22'
22'	4.09 (1H, d, 18.5)	60.0	15, 16, 17, 18	15, 17, 22
-COOH	12.01 (1H, br. s)	-	-	-

Yanuthone X₁

HRESIMS: $m/z = 403.2482$ $[M + H]^+$, calculated for $[C_{17}H_{24}O_6 + H]^+$: 403.2479 $[\alpha]_{587}^{20} = 2.5^\circ$

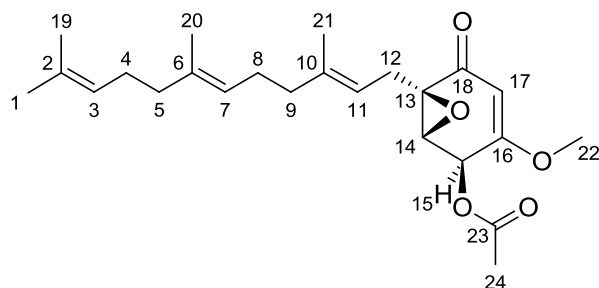


Table S12. NMR data for yanuthone X₁

Atom assignment	¹ H-chemical shift [ppm]/ J coupling constants [Hz]	¹³ C-chemical shift [ppm]	HMBC correlations	NOESY correlations
1	1.66 (3H, s)	25.6	2, 3, 19	3
2	-	132.2	-	-
3	5.08 (1H, m)	125.0	5	1, 4/5
4	2.05 (2H, m)	27.6	3, 5, 6	3, 5
5	2.00 (2H, m)	40.3	6, 4/8, 7, 11	3, 4
6	-	135.6	-	-
7	5.08 (1H, m)	125.0	5/9, 20	-
8	2.07 (2H, m)	27.6	5/9, 6, 7	9
9	1.95 (2H, m)	40.3	7, 8, 10, 11, 21	8, 11
10	-	139.8	-	-
11	5.01 (1H, t, 7.3)	117.5	9, 21	9, 12, 12'
12	2.78 (1H, dd, 15.3, 8.2)	26.2	10, 11, 13, 14	11, 12'
12'	2.45 (1H, dd, 15.3, 6.7)	26.2	10, 11, 13, 14, 18	11, 12
13	-	60.5	-	-
14	3.59 (1H, d, 3.1)	56.3	15, 16	15
15	5.95 (1H, d, 3.1)	66.6	16	14
16	-	168.3	-	-
17	5.30 (1H, s)	100.3	13, 15, 16	22
18	-	193.7	-	-
19	1.57 (3H, s)	17.6	1, 2, 3	-
20	1.571 (3H, s)	15.8	5, 6	-
21	1.61 (3H, s)	16.3	9, 10, 11	-
22	3.67 (3H, s)	57.4	16	17
23	-	170.7	-	-
24	2.14 (3H, s)	20.6	23	-

Table S13. Fungal strains.

Name	Organism	Genotype
IBT 29539	<i>A. nidulans</i>	<i>argB2, pyrG89, veA1, nkuAΔ</i>
OE- <i>yanA</i>	<i>A. nidulans</i>	<i>argB2, pyrG89, veA1, nkuAΔ S1::PgpA::yanA::TtrpC::argB</i>
OE- <i>yanG</i>	<i>A. nidulans</i>	<i>argB2, pyrG89, veA1, nkuAΔ IS1::PgpA::yanA::TtrpC::argB, pDHX2::yanG</i>
OE- <i>yanC</i>	<i>A. nidulans</i>	<i>argB2, pyrG89, veA1, nkuAΔ IS1::PgpA::yanA::TtrpC::argB, pDHX2::yanC</i>
OE- <i>yanB</i>	<i>A. nidulans</i>	<i>argB2, pyrG89, veA1, nkuAΔ IS1::PgpA::yanA::TtrpC::argB, pDHX2::yanB</i>
OE- <i>yanD</i>	<i>A. nidulans</i>	<i>argB2, pyrG89, veA1, nkuAΔ IS1::PgpA::yanA::TtrpC::argB, pDHX2::yanD</i>
OE- <i>yanE</i>	<i>A. nidulans</i>	<i>argB2, pyrG89, veA1, nkuAΔ IS1::PgpA::yanA::TtrpC::argB, pDHX2::yanE</i>
KB1001	<i>A. niger</i>	<i>pyrGΔ kusA::AFpyrG</i>
<i>yanAΔ</i>	<i>A. niger</i>	<i>pyrGΔ kusA::AFpyrG yanAΔ</i>
<i>yanGΔ</i>	<i>A. niger</i>	<i>pyrGΔ kusA::AFpyrG yanGΔ</i>
<i>yanIΔ</i>	<i>A. niger</i>	<i>pyrGΔ kusA::AFpyrG yanIΔ</i>
<i>yanCΔ</i>	<i>A. niger</i>	<i>pyrGΔ kusA::AFpyrG yanCΔ</i>
<i>yanBΔ</i>	<i>A. niger</i>	<i>pyrGΔ kusA::AFpyrG yanBΔ</i>
<i>yanHΔ</i>	<i>A. niger</i>	<i>pyrGΔ kusA::AFpyrG yanHΔ</i>
<i>yanDΔ</i>	<i>A. niger</i>	<i>pyrGΔ kusA::AFpyrG yanDΔ</i>
<i>yanEΔ</i>	<i>A. niger</i>	<i>pyrGΔ kusA::AFpyrG yanEΔ</i>
<i>yanFΔ</i>	<i>A. niger</i>	<i>pyrGΔ kusA::AFpyrG yanFΔ</i>
<i>yanRΔ</i>	<i>A. niger</i>	<i>pyrGΔ kusA::AFpyrG yanRΔ</i>
44959Δ	<i>A. niger</i>	<i>pyrGΔ kusA::AFpyrG ASPNI_DRAFT44959Δ</i>
44960Δ	<i>A. niger</i>	<i>pyrGΔ kusA::AFpyrG ASPNI_DRAFT44960Δ</i>
44971Δ	<i>A. niger</i>	<i>pyrGΔ kusA::AFpyrG ASPNI_DRAFT44971Δ</i>
44972Δ	<i>A. niger</i>	<i>pyrGΔ kusA::AFpyrG ASPNI_DRAFT44972Δ</i>
ClusterΔ	<i>A. niger</i>	<i>pyrGΔ kusA::AFpyrG ASPNI_DRAFT44958-44972Δ</i>

Table S14. Primers used in the study. See details in experimental section.

#	Primer name	Sequence 5'→ 3'
1	44965-fw	AGAGCGAUATGCCAGGCCTTGTACAC
2	44965-rv	TCTGCGAUTTAAGCATCCAGCTCCTTTGT
3	44963_ORF_FW	AGAGCGAUATGGACCGTATCGACGTACACC
4	44963_ORF_RV	TCTGCGAUCTAGGTACTATAAGTATGAACACGAGACTG
5	44964_ORF_FW	AGAGCGAUATGTCTACTACTAAGCGCTCGGTAAC
6	44964_ORF_RV	TCTGCGAUCTAGTATACTTTTCATGGGTGCGTGA
7	54844_ORF_FW	AGAGCGAUCGGGCTAGACTTTTCTCTTCTTAAG
8	54844_ORF_RV	TCTGCGAUATGGCGCTTGTTTCATCTGACT
9	127904_ORF_FW	AGAGCGAUATGGTCAAGTTTTTTCAGCCCA
10	127904_ORF_RV	TCTGCGAUCTAACGGAAGTGGGGAGGAA
11	192604_ORF_FW	AGAGCGAUTACTTTGCGACTACCTGCCATG
12	192604_ORF_RV	TCTGCGAUCTACTCCGACTTTTCACCTTTGG
13	hph-1003-Fw	AGCCCAATAUGCTAGTGGAGGTCAACACATCA
14	hph-1003-Rv	ATTACCTAGUCGGTCGGCATCTACTCTATT
15	44965-chk-usF	CAGTTGACTAGACTAGGAACGGTCA
16	44965-chk-dsR	AACGACCATGATGCTTGTTTCAG
17	44965_US-FW	GGGTTTAAUATGACTCCACATCATCTTCCACAC
18	44965_US-RV	ATATTGGGCUGATGGTGTGTACAAGGCCTGG
19	44965_DS-FW	ACTAGGTAAUGACTGTTATGCATTGAATTTGAGC
20	44965_DS-RV	GGTCTTAAUAGATCCTGACGCTCATATCTGCT
21	44964_US-FW	GGGTTTAAUGGTCTTTCCGACACGTAAGTCTG
22	44964_US-RV	ATATTGGGCUTGGACCTCAATGGCCGCT
23	44964_DS-FW	ACTAGGTAAUGCGAGTATGAAGAAGGTGGATGA
24	44964_DS-RV	GGTCTTAAUATTCAGGGTCTTGAGATTGGC
25	44963_US-FW	GGGTTTAAUAAGTCCTCCACGTCGGAG
26	44963_US-RV	ATATTGGGCUAGAATCTAAACCTTGTCTCTTCGCT
27	44963_DS-FW	ACTAGGTAAUGAACGTTTGATTGGTAATGGATGT
28	44963_DS-RV	GGTCTTAAUTCATCCACCTTCTTCATACTCGC
29	54844_US-FW	GGGTTTAAUGTAGAATAACAGCTACCTCGAATTTGA
30	54844_US-RV	ATATTGGGCUACGTGGTGCGTAAGCAGACAT
31	54844_DS-FW	ACTAGGTAAUCCTGCTGAATAAACACGAAGG
32	54844_DS-RV	GGTCTTAAUATGGACCGTATCGACGTACACC
33	44960_US-FW	GGGTTTAAUTGAGTACCTATCCACTCTTCTCTGG
34	44960_US-RV	ATATTGGGCUGATGGAGTGTGAAGCCAATGAG
35	44960_DS-FW	ACTAGGTAAUTCATTCTAAAATTGGCGTCTTCA
36	44960_DS-RV	GGTCTTAAUCTACTGCCGCCGTCACTATCTA
37	193092_US-FW	GGGTTTAAUCATCGACATCTCTCTGCCCAT

38	193092_US-RV	ATATTGGGCUGAAAGCTGGTTGGAAGTATAAGTGG
39	193092_DS-FW	ACTAGGTAAUTGTGCAGCGGTATTGACTTCA
40	193092_DS-RV	GGTCTTAAUCACGGAGTTATTTTCCACGCT
41	44967_US-FW	GGGTTTAAUCGTTGGCATGACAGTCTTCAA
42	44967_US-RV	ATATTGGGCUGTCTGCCATCACAACCAGTTTG
43	44967_DS-FW	ACTAGGTAAUAGCCATGTTGCCAGACACAGT
44	44967_DS-RV	GGTCTTAAUACTACCATCTCGTAACCGTCCTAG
45	127904_US-FW	GGGTTTAAUGACCGACTCTACACTACCGTTCC
46	127904_US-RV	ATATTGGGCUATTGAACTGGTAAACATGCCATG
47	127904_DS-FW	ACTAGGTAAUTAGCCCTAGGACGGTTACGAG
48	127904_DS-RV	GGTCTTAAUAACCAACTTTGTTCCATTCTATCG
49	192604_US-FW	GGGTTTAAUGACACATCGTATTGATGACGACC
50	192604_US-RV	ATATTGGGCUCATGGCAGGTAGTCGCAAAG
51	192604_DS-FW	ACTAGGTAAUAAGAGAATACGGAACACATTGACC
52	192604_DS-RV	GGTCTTAAUCGGTCCAACAGTGAGGGTCT
53	44970_US-FW	GGGTTTAAUCGTTGATAATTCCAATTCCAATTC
54	44970_US-RV	ATATTGGGCUCGTGCAAGATGACCTGATTTG
55	44970_DS-FW	ACTAGGTAAUCGGGTATCACTGTATCAATATCG
56	44970_DS-RV	GGTCTTAAUGCTACTACTATGCCGACTGCGT
57	44971_US-FW	GGGTTTAAUGGCCACACCTCAAGTTTGTATG
58	44971_US-RV	ATATTGGGCUCGGGATTGGAGTGCTCTAGTT
59	44971_DS-FW	ACTAGGTAAUGTTGGCTGAGAGTCAGGGTTAG
60	44971_DS-RV	GGTCTTAAUCCATTAGCTTCGGAACACTGG
61	44959_US-FW	GGGTTTAAUCCTTGTATTCATATCAATTGCGA
62	44959_US-RV	ATATTGGGCUATGTGACAATGAAGAATGGTACG
63	44959_DS-FW	ACTAGGTAAUGGAAAGGATGTTCCAAACAGTT
64	44959_DS-RV	GGTCTTAAUCTTTGTTGATTACTAGTCGTAATCATATG
65	44961_US-FW	GGGTTTAAUTGTCATGTTGTATCGGAGTGTTTAG
66	44961_US-RV	ATATTGGGCUTGTAGCACAAGTGTCTCACTAGTAAATAG
67	44961_DS-FW	ACTAGGTAAUGATTGGAAGTATCCCACAGTCTG
68	44961_DS-RV	GGTCTTAAUGAGAACACCGATCTCCGACGTGGGA
69	44972_US-FW	GGGTTTAAUGACGCAGTCGGCATAGTAGTAG
70	44972_US-RV	ATATTGGGCUGGAGAAGTGGTCAAACCTTGTTTCA
71	44972_DS-FW	ACTAGGTAAUACAGGTGATTAAGATGCAAGGCT
72	44972_DS-RV	GGTCTTAAUCTTGCATCATCCGTAATTATGCT
73	Upst-HygR-N	CTGCTGCTCCATACAAGCCAACC
74	Dwst-1003HygF-N	GACATTGGGGAGTTCAGCGAGAG
75	ampR_PM_FW	AGCGCTACAUAATTCTCTTACTGTCATGCCATCC
76	ampR_PM_RV	ATGTAGCGCUGCCATAACCATGAGTGATAACACTG
77	ori_coli_FW	ATCCCCACUACCGCATTAAGACCTCAGCG

78	ampR_RV	AGCTGCTUCGTCGATTAAACCCTCAGCG
79	p71_prom- ter_short_usF	AGCCCAATAUTAAGCTCCCTAATTGGCCC
80	p71_prom-ter_dsR	ATTACCTAGUGGGCGCTTACACAGTACA
81	argB_FW	ACTAGGTAAUATCGCGTGCATTCCGCGGT
82	pyrG_FW_C1	ACTAGGTAAUATGACATGATTACGAATTCGAGCT
83	pyrG_RV_G16	AGTGGGGAUGCCTCAATTGTGCTAGCTGC
84	ccdB-camR-fw	AGAGCGAUCGCAGAAGCCTACTCGCTATTGTCCTCA
85	ccdB-camR-rv	TCTGCGAUCGCTCTTGCGCCGAATAAATACCTGT
86	AMA1_FW	AAGCAGCUGACGGCCAGTGCCAAGCT
87	AMA1_RV	ATATTGGGCUGGAAACAGCTATGACCATGAGATCT
88	AMA-3'-Fw	ACCCCAAUGGAAACGGTGAGAGTCCAGTG
89	AMA-5'-Rv	ATTGGGGUACTAACATAGCCATCAAATGCC
90	ANIG-actA-qFw	GTATGCAGAAGGAGATCACTGCTCT
91	ANIG-actA-qRv	GAGGGACCGCTCTCGTCGT
92	ANIG-hhtA-qFw	CTTCCAGCGTCTTGTCCGTG
93	ANIG-hhtA-qRv	GCTGGATGTCCTTGGACTGGAT
94	44959-qFw	GGCAAAGTTCTAGTCATCGACGA
95	44959-qRv	CATATATCCCAGAGGGCGGACAC
96	44960-qFw	GATAGAGGAGATGAGGAAGAGAGGCT
97	44960-qRv	CCTTGGGTACCATTACAGTCAG
98	44961-qFw	ACATGGACCACCGAGTAGCGT
99	44961-qRv	TAGGGTGTGCGAGAATATCACTTG
100	54844-qFw	CGATGAAGATGGCAATCCCAT
101	54844-qRv	CTATGGCATCGCATACTGAGAAAGA
102	44963-qFw	GCGAGGAGGTAGAAAAGGCAAT
103	44963-qRv	TGAACACGAGACTGGAGTACGGA
104	44964-qFw	TCTTCTGGATACTGGGAATTGGAG
105	44964-qRv	GCGTGATGCACCCTCAACA
106	44965-qFw	TTGTCTGTCAAGGAGGACGAGATT
107	44965-qRv	CTTCACCAAATGCTGCACAGTC
108	193092-qFw	ATCACGGCAAAGAGAGCCAAGT
109	193092-qRv	GAAGTGTGGCACGACCATGTC
110	44967-qFw	TGCTGGCTGCTAAGGATTGATG
111	44967-qRv	ATGTTCCAACGCAATGAACAAC
112	127904-qFw	ATGAGCAACATGCTCCCACTACAT
113	127904-qRv	CAGTGATGGTCTTATCCGCCAG
114	192604-qFw	GCCTAATCCTGGGCATCGTG
115	192604-qRv	CTGTGCTCCCCGATCTGCA
116	44970-qFw	TCTCAGGGTGTCCATCTTCCGT

117	44970-qRv	CGACACGAAATAGGCATCATTCT
118	44971-qFw	ATCTACTCCGGCTCCTGCGAT
119	44971-qRv	ACTCGCAAACAAC TTCATTGCTC
120	44972-qFw	AGGTGACTCGAACTGGTATGCTG
121	44972-qRv	CAGAATATACTCGATATGATCGCCTC

Table S15. ^{13}C NMR chemical shift in ppm for compound 1-9 and 12.

Atom #	1	2	4	5	6	7	8	9	12
1	17.1	17.7	68.3	68.2	25.8	68.3	29.7		25.6
2	132.1	131.8	136.2	135.9	132.0	136.0	70.9		132.2
3	125.0	124.9	125.3	125.2	124.9	125.4	44.0		125.0
4	27.3	27.2	26.8	26.8	26.8	26.8	23.1		27.6
5	40.2	40.3	40.1	40.0	40.1	40.1	40.5	177.4	40.3
6	136.4	135.8	136.2	135.9	136.0	135.9	136.3	38.2	135.6
7	124.9	124.9	124.9	124.9	124.9	124.9	124.4	32.5	125.0
8	27.0	27.2	26.8	26.9	26.8	26.8	26.7	24.5	27.6
9	40.1	40.3	40.1	40.0	40.1	40.1	40.3	38.7	40.3
10	141.0	139.9	139.7	139.7	140.2	139.8	139.6	137.9	139.8
11	116.4	117.7	118.0	117.9	118.0	118.0	117.5	116.9	117.5
12	25.6	26.6	27.0	26.9	26.8	26.8	26.5	25.7	26.2
13	63.0	61.3	61.4	61.3	61.6	60.9	61.1	60.0	60.5
14	58.6	60.1	60.2	59.7	60.2	60.4	59.9	59.1	56.3
15	193.2	65.8	67.6	76.2	66.2	66.1	65.7	64.3	66.6
16	143.9	154.6	158.8	157.0	160.8	160.6	154.7	162.2	168.3
17	133.1	121.6	123.2	123.9	119.7	119.7	121.5	117.4	100.3
18	193.2	194.8	194.9	194.4	195.1	194.9	194.2	193.7	193.7
19	25.7	25.6	13.5	13.7	17.7	13.5	29.2		17.6
20	16.1	16.0	16.1	15.9	15.9	16.1	15.9	16.7	15.8
21	16.2	16.3	16.1	16.3	16.2	16.2	16.2	15.6	16.3
22	60.2	63.6	20.0	20.1	62.0	62.0	63.6	60.0	57.4
23	171.2	171.4		105.7			171.2		170.7
24	45.6	45.8		74.6			43.7		20.6
25	70.0	70.2		77.2			70.9		
26	45.1	45.4		71.2			45.1		
27	173.1	173.5		77.2			173.6		
28	27.6	27.7		62.6			27.5		

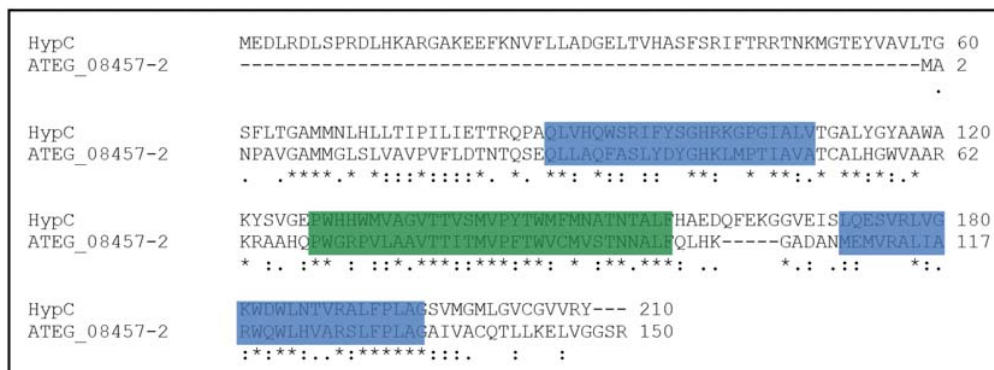
Table S16. ¹H NMR chemical shift in ppm for compound 1-9 and 12.

Atom #	1	2	4	5	6	7	8	9	12
1	1.59 (3H, s)	1.58 (3H, s)	3.84 (2H, br. s)	3.86 (2H, s)	1.66 (3H, s)	3.85 (2H, s)	1.28 (3H, m)	-	1.66 (3H, s)
2	-	-	-	-	-	-	-	-	-
3	5.10 (1H, m)	5.07 (1H, m)	5.33 (1H, tq, 7.4, 1.5)	5.34 (1H, t, 6.3)	5.10 (1H, m)	5.34 (1H, m)	1.35 (2H, m)	-	5.08 (1H, m)
4	2.06 (2H, m)	2.02 (2H, m)	2.11 (2H, m)	2.10 (2H, m)	2.07 (2H, m)	2.11 (2H, m)	1.43 (2H, m)	-	2.05 (2H, m)
5	1.97 (2H, m)	1.96 (2H, m)	1.99 (2H, m)	1.99 (2H, m)	1.98 (2H, m)	1.99 (2H, m)	1.93 (2H, m)	-	2.00 (2H, m)
6	-	-	-	-	-	-	-	2.28 (1H, m)	-
7	5.09 (1H, m)	5.06 (1H, t, 6.4)	5.09 (1H, tq, 6.9, 1.1)	5.10 (1H, t, 6.3)	5.10 (1H, m)	5.10 (1H, t, 6.4)	5.08 (1H, tq, 6.3, 1.0)	1.48 (1H, m)	5.08 (1H, m)
7'	-	-	-	-	-	-	-	1.26 (1H, m)	-
8	2.09 (2H, m)	2.06 (2H, m)	2.07 (2H, m)	2.06 (2H, m)	2.09 (2H, m)	2.07 (2H, m)	2.09 (2H, m)	1.32 (2H, m)	2.07 (2H, m)
9	1.99 (2H, m)	1.99 (2H, m)	2.01 (2H, m)	2.02 (2H, m)	2.03 (2H, m)	2.02 (2H, m)	2.02 (2H, m)	1.92 (2H, t, 7.1)	1.95 (2H, m)
10	-	-	-	-	-	-	-	-	-
11	5.03 (1H, t, 1.5)	5.04 (1H, t, 7.2)	5.04 (1H, tq, 7.3, 1.2)	5.04 (1H, t, 6.7)	5.05 (1H, ddd, 7.93, 6.71, 1.22)	5.05 (1H, t, 6.8)	5.03 (1H, tq, 6.3, 1.0)	5.00 (1H, t, 7.1)	5.01 (1H, t, 7.3)
12	2.66 (1H, m)	2.67 (1H, m)	2.70 (1H, dd, 15.1, 8.1)	2.67 (1H, m)	2.70 (1H, dd, 15.3, 7.9)	2.71 (1H, dd, 15.1, 8.3)	2.45 (1H, m)	2.63 (1H, dd, 15.1, 7.9)	2.78 (1H, dd, 15.3, 8.2)
12'	2.76 (1H, m)	2.42 (1H, m)	2.37 (1H, m)	2.37 (1H, dd, 14.8, 6.2)	2.42 (1H, dd, 15.3, 6.7)	2.40 (1H, dd, 15.1, 6.8)	2.68 (1H, m)	2.32 (1H, m)	2.45 (1H, dd, 15.3, 6.7)
13	-	-	-	-	-	-	-	-	-
14	3.68 (1H, s)	3.64 (1H, d, 2.8)	3.60 (1H, d, 2.7)	3.82 (1H, br. s)	3.62 (1H, d, 2.8)	3.62 (1H, d, 2.9)	3.63 (1H, d, 2.8)	3.59 (1H, d, 2.4)	3.59 (1H, d, 3.1)
15	-	4.68 (1H, br. s)	4.48 (1H, m)	4.59 (1H, m)	4.64 (1H, m)	4.64 (1H, br. s)	4.69 (1H, m)	4.61 (1H, br. S)	5.95 (1H, d, 3.1)
16	-	-	-	-	-	-	-	-	-
17	6.58 (1H, t, 1.5)	5.86 (1H, q, 1.4)	5.68 (1H, m)	5.71 (1H, br. s)	5.87 (1H, q, 1.73)	5.87 (1H, d, 1.5)	5.86 (1H, q, 1.5)	5.82 (1H, d, 1.6)	5.30 (1H, s)
18	-	-	-	-	-	-	-	-	-
19	1.66 (3H, s)	1.66 (3H, s)	1.59 (3H, s)	1.591 (3H, s)	1.601 (3H, s)	1.601 (3H, s)	1.12 (3H, s)	-	1.57 (3H, s)
20	1.591 (3H, s)	1.581 (3H, s)	1.591 (3H, s)	1.59 (3H, s)	1.60 (3H, s)	1.60 (3H, s)	1.58 (3H, s)	1.02 (3H, d, 6.7)	1.571 (3H, s)
21	1.63 (3H, s)	1.62 (3H, s)	1.62 (3H, s)	1.62 (3H, s)	1.63 (3H, s)	1.63 (3H, s)	1.62 (3H, s)	1.56 (3H, s)	1.61 (3H, s)
22	4.84 (1H, dd, 16.1, 1.5)	4.82 (1H, d, 16.1)	1.92 (2H, br. s)	1.97 (3H, m)	4.27 (1H, m)	4.27 (1H, d, 17.6)	4.83 (1H, d, 16.0)	4.21 (1H, d, 18.5)	3.67 (3H, s)
22'	4.89 (1H, dd, 16.1, 1.5)	4.76 (1H, d, 16.1)	-	-	4.21 (1H, m)	4.21 (1H, m)	4.76 (1H, d, 16.0)	4.09 (1H, d, 18.5)	-
23	-	-	-	4.56 (1H, m)	-	-	-	-	-
24	2.69 (2H, m)	2.69 (2H, m)	-	3.25 (1H, br. s)	-	-	2.681 (2H, m)	-	2.14 (3H, s)
25	-	-	-	3.37 (1H, m)	-	-	-	-	-
26	2.64 (1H, m)	2.59 (2H, m)	-	3.35 (1H, m)	-	-	2.61 (2H, m)	-	-
26'	2.57 (1H, m)	-	-	-	-	-	-	-	-
27	-	-	-	3.35 (1H, m)	-	-	-	-	-
28	1.33 (3H, s)	1.33 (3H, s)	-	6.77 (1H, d, 10.7)	-	-	1.33 (3H, s)	-	-
28'	-	-	-	3.66 (1H, d, 10.7)	-	-	-	-	-

Appendix F

Supporting information for Chapter 11

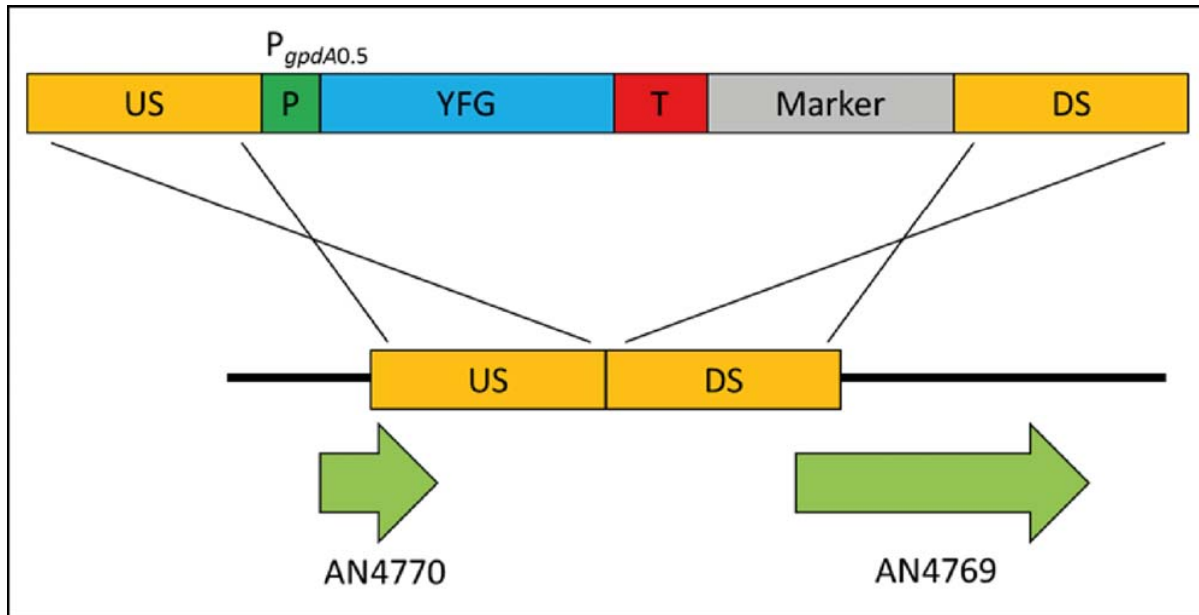
Figure S1.



Identification of putative HypC homolog encoded by *gedH* (ATEG_08457-2) in the *A. terreus* geodin gene cluster. Pairwise alignment of putative emodin anthrone oxidase, GedH (ATEG_08457-2), from *A. terreus* and norsolic anthrone oxidase, HypC, from *A. flavus*. The conserved DUF-1772 domain and putative catalytic regions proposed by Ehrlich et al [25] are highlighted in green and blue, respectively.

doi:10.1371/journal.pone.0072871.s001

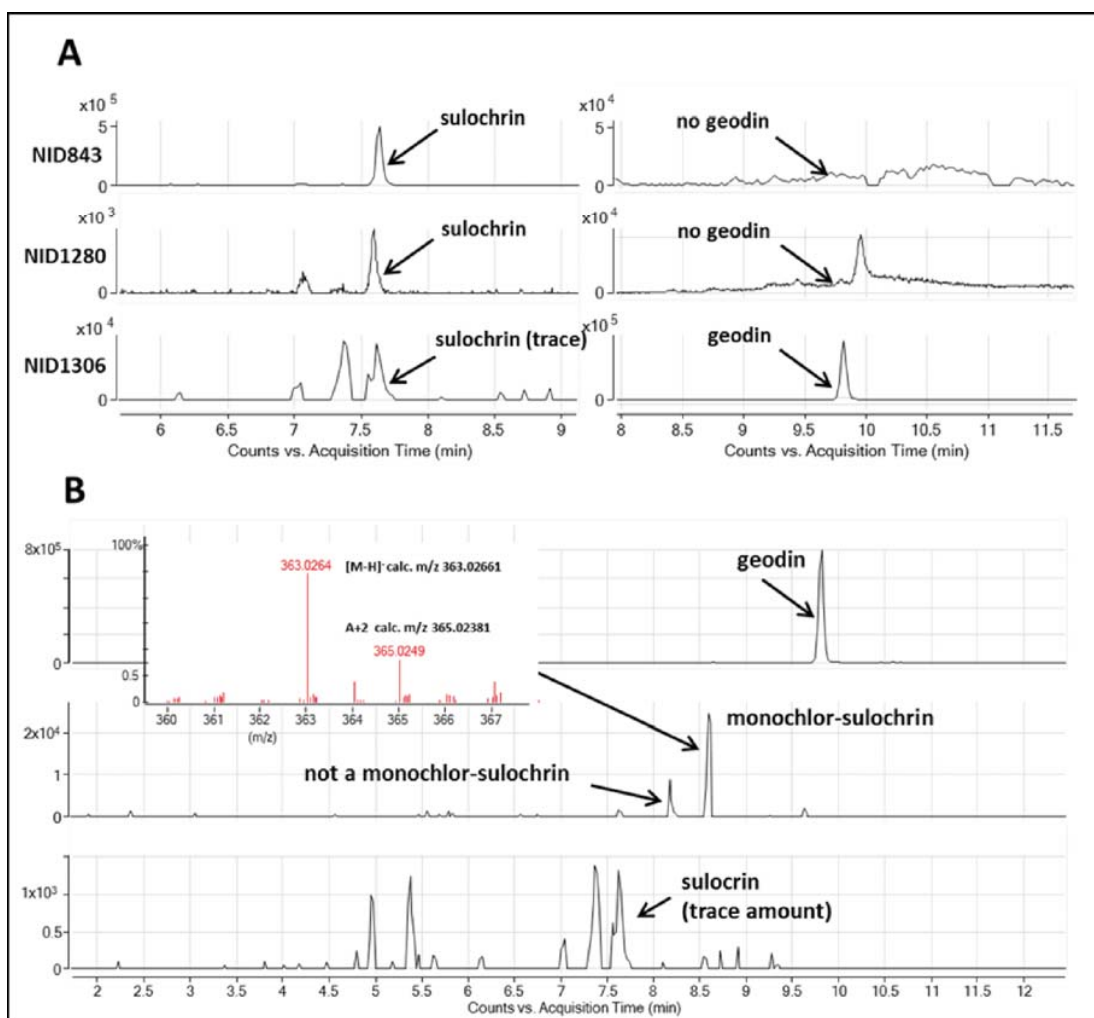
Figure S2.



Schematic overview of the integration of a gene-expression cassette into the integration site, *IS3*, by homologous recombination. *IS3* is located between genes AN4770 and AN4769 on chromosome III. The cassette consists of six parts: upstream targeting sequence (US), promoter (P, in this case 0.5 kb *P_{gpdA}*), your favorite gene (YFG), terminator (T, *T_{trpC}*), marker (in this case *AF_{pyrG}* flanked by direct), and the downstream targeting sequence (DS). The orientations of the genes AN4770 and AN4769 are indicated by green arrows. The sizes of US, DS and the intergenic region are 1984 bp, 1911 bp, and 3007 bp, respectively.

doi:10.1371/journal.pone.0072871.s002

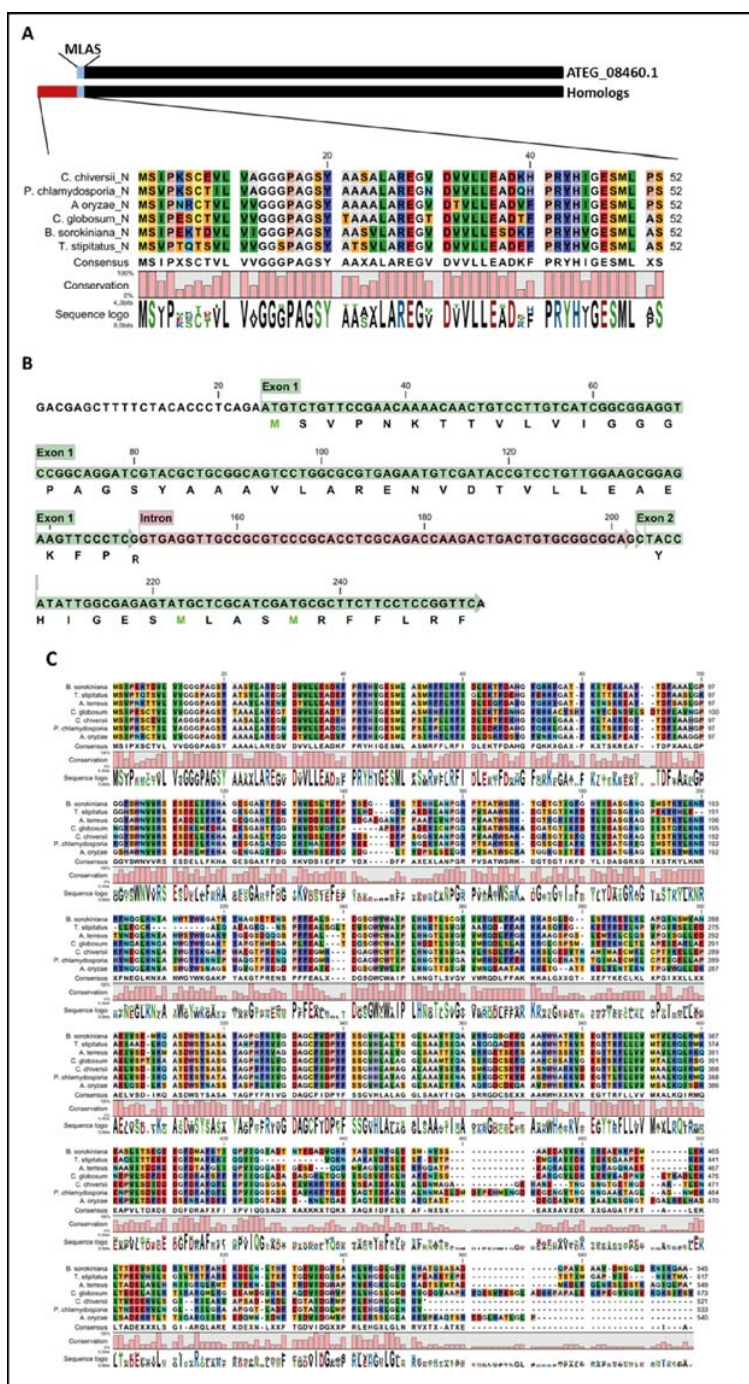
Figure S3.



Complementation of halogenase deficiency. A) Left panel: detection of sulochrin (-ESI, EIC(m/z 331.0812)); right panel: detection of geodin (-ESI, EIC(m/z 396.9876)). Strains for halogenase analysis, from top to bottom: NID843 (*gedL* Δ); NID1280 (*gedL* Δ , *IS3::PgpA-ATEG_08460.1*); and NID1306 (*gedL* Δ , *IS2::gedL*). **B)** Ratio of geodin, monochlor-sulochrin, and sulochrin in the NID1306 strain, including ESI⁻-MS spectrum of monochlor-sulochrin showing the isotopic pattern and the mass deviations relative to the theoretical masses. Reference standards of geodin and sulochrin were included in all runs (data not shown).

doi:10.1371/journal.pone.0072871.s003

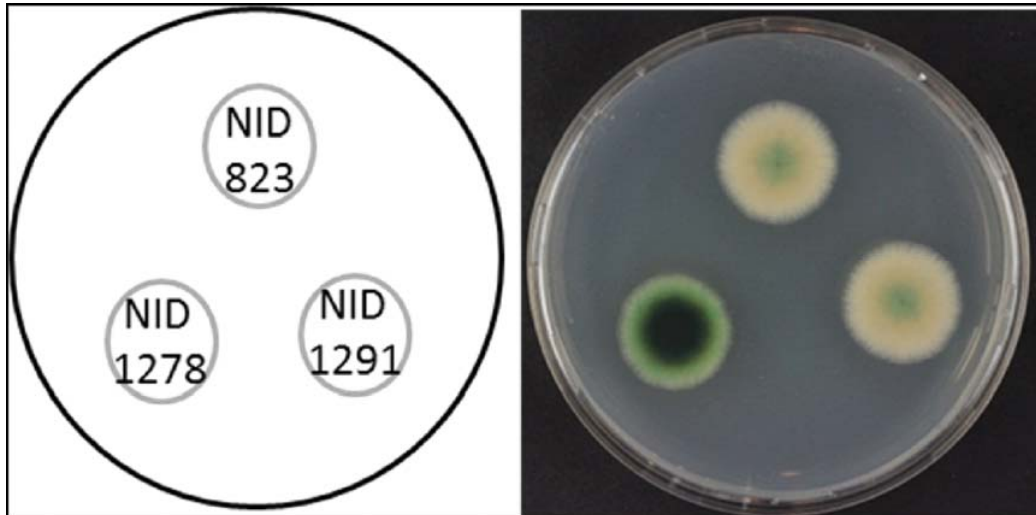
Figure S4.



Identification of the likely start codon of *gedL*. **A)** Alignment of the top hits in a BLAST search for ATEG_08460.1 homologs shows that they contain a very conserved 48 amino acid residue addition in the N-terminus. Amongst the homologs, Rdc2, has been characterized as a halogenase by [39] MLAS is the predicted N-terminus of ATEG_08460.1. Drawing is not to scale. **B)** The position of putative exons and intron in the 5' end of *gedL* as predicted by the Augustus software [29]. The predicted protein sequence encoded by exon 1 and by the first section of exon 2 is indicated. **C)** Full alignment of the halogenase homologs and GedL based on the GedL sequence derived from the new start codon.

doi:10.1371/journal.pone.0072871.s004

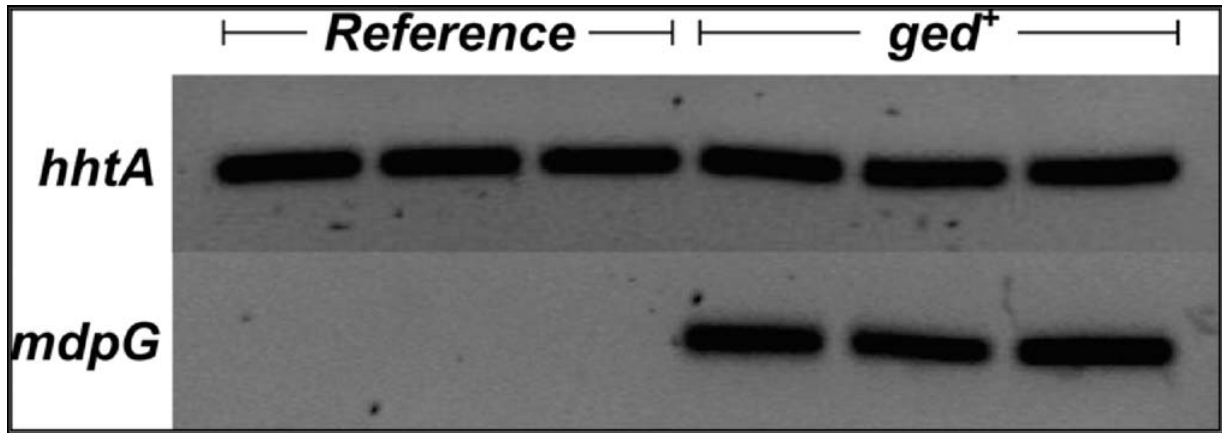
Figure S5.



Expression of *lacZ* under the control of the *gedR* promoter (*PgedR*). Left panel: the positions of the strains on the plate are shown in the right panel. NID823 (*ged*⁺ *mdpA-LΔ*) is the reference strain without the *lacZ* gene. NID1278 is a control strain containing the *PgpdA-lacZ* construct in *IS3*. The NID1291 (*ged*⁺ *mdpA-LΔ PgedR-lacZ*) strain carries *PgedR-lacZ* in *IS3*. The strains were stabbed on MM containing X-gal and incubated three days at 37 °C in the dark before photography.

doi:10.1371/journal.pone.0072871.s005

Figure S6.



Constitutive expression of *gedR* induces transcription of the *A. nidulans* gene *mdpG*. *mdpG* mRNA levels in reference (NID1) and in the *ged⁺* strain (NID677) were evaluated by quantitative RT-PCR. For each strain, RNA was extracted as described in [Materials and Method](#) and the RNA samples analyzed in triplicate by quantitative RT-PCR. The samples were loaded and analyzed by 1% agarose gel-electrophoresis as indicated in the figure.

doi:10.1371/journal.pone.0072871.s006

Table S1.

Strain genotypes. * = For reference see Nielsen et al ([13]).

doi:10.1371/journal.pone.0072871.s007

Strain	Genotype
NID1	<i>argB2, pyrG89, veA1, nkuAΔ</i>
NID3	<i>argB2, pyrG89, veA1, nkuA-trS::AFpyrG</i>
NID74	<i>argBΔ, pyrG89, veA1, nkuAΔ</i>
NID356	<i>argBΔ, pyrG89, veA1, nkuAΔ, mdpA-LΔ</i>
NID675	<i>argBΔ, pyrG89, veA1, nkuAΔ, IS1::ATgedE-L::argB</i>
NID676	<i>argBΔ, pyrG89, veA1, nkuAΔ, mdpA-LΔ, IS1::ATgedE-L::argB</i>
NID677	<i>argBΔ, pyrG89, veA1, nkuAΔ, IS1::ATgedA-L::PgpdA-gedR::AFpyrG</i>
NID695	<i>argBΔ, pyrG89, veA1, nkuAΔ, mdpA-LΔ, IS1::ATgedA-L::PgpdA-gedR::AFpyrG</i>
NID761	<i>argBΔ, pyrG89, veA1, nkuAΔ, mdpA-LΔ IS1::ATgedA-L, gedRΔ:AFpyrG</i>
NID802	<i>argBΔ, pyrG89, veA1, nkuAΔ, IS1::ATgedA-L::PgpdA-gedR</i>
NID823	<i>argBΔ, pyrG89, veA1, nkuAΔ, mdpA-LΔ, IS1::ATgedA-L::PgpdA-gedR</i>
NID842	<i>argBΔ, pyrG89, veA1, nkuAΔ, mdpA-LΔ, IS1::ATgedA-B::gedD-gedL::PgpdA-gedR, gedCΔ:AFpyrG</i>
NID843	<i>argBΔ, pyrG89, veA1, nkuAΔ, mdpA-LΔ, IS1::ATgedA-K::PgpdA-gedR, gedLΔ:AFpyrG</i>
NID925	<i>argBΔ, pyrG89, veA1, nkuAΔ, mdpA-LΔ, IS1::ATgedA-L, gedRΔ</i>
NID1278	<i>argB2, pyrG89, veA1, nkuAΔ, IS3::PgpdA-lacZ</i>
NID1279	<i>argBΔ, pyrG89, veA1, nkuAΔ, mdpA-LΔ, IS1::ATgedA-K::PgpdA-gedR, gedLΔ</i>
NID1280	<i>argB2, pyrG89, veA1, nkuAΔ, IS1::ATgedA-K::PgpdA-gedR, gedLΔ, IS3::PgpdA-ATEG_08460.1::AFpyrG</i>
NID1291	<i>argBΔ, pyrG89, veA1, nkuAΔ, mdpA-LΔ IS1::ATgedA-L, IS3::0.9 kb PgedR-lacZ</i>
NID1297	<i>argBΔ, pyrG89, veA1, nkuAΔ, mdpA-LΔ IS1::ATgedA-L, gedRΔ, IS3::0.9 kb PgedR-lacZ</i>
NID1306	<i>argBΔ, pyrG89, veA1, nkuAΔ, mdpA-LΔ, IS1::ATgedA-K::PgpdA-gedR, gedLΔ:AFpyrG, IS2::gedL</i>

Table S2.

Primers list. The sequence of PCR products # 1-3 corresponds to supercontig 12: 1302695-1315163 (ATEG_08454-08460) in the genome sequenced isolate NIH2624 (IBT28053). Similarly, PCR products # 4-6 and 8 corresponds to supercontig 12: 1289727-1304484 (ATEG_08449-08454) in NIH2624 (IBT28053). * = For reference see Hansen et al ([7]).

doi:10.1371/journal.pone.0072871.s008

PCR product #	Primer pair (fw/rv)	Forward primer	Reverse Primer	Description
1	455/456	agctcacuCCAGATCA ACCAGCAGACGA	aggcccagguGACGTTT GCGATTCCGCCCA	4.072 bp of <i>A.terreus</i> IBT15722 gDNA spanning putative ORFs ATEG_8454-8456 including 1788 bp overlap with 521/464 PCR product
2	457/177	acctgggcuCGTGAG TTGGGATGGAATG A	aacacaauGAAGGCTG TGTGATGCCCGG	4.382 bp of <i>A.terreus</i> IBT15722 gDNA spanning putative ORFs ATEG_8457-8458
3	178/508	attgtgtuAGATAACG CCAGAACTGCCTA	actggagauCGAGGAT ATGGAAGGACACTA	3.993 bp of <i>A.terreus</i> IBT15722 gDNA spanning putative ORFs ATEG_8459-8460
4	459/460	agctcacuTTAGAACC CAGCAGCAAGGCG	agcggacacuCCTGTCG CTGCAGTTCAGTT	4.430 bp of <i>A.terreus</i> IBT15722 gDNA spanning putative ORFs ATEG_8449-8451 (5')
5	461/462	agtgtccguAACGCT TAGAGGCAGCTACT	atctgcagacuGGATGA CTAGATGGCACTGG	5.382 bp of <i>A.terreus</i> IBT15722 gDNA spanning the 3' end of the Geodin PKS (ATEG_8451)
6	463/522	agtctgcagauAATGA GGTTACAATGCATC C	agcttacuTGTGATGCT CCCTGGTGACT	2.397 bp of <i>A.terreus</i> IBT15722 gDNA spanning putative ORF ATEG_8452
7	405/406	agtaagcuCCCTAATT GGCCCATCCG	atctgaaaaGCGGTAG TGATGTCTGCTCA	500 bp of <i>A. nidulans</i> PgpA promoter amplified from pU1111 (20)
8	521/464	attttcagauGTCTGC AGAAGGACACAAG C	actggagauCATGGATT TGCGGCAGTACT	2.533 bp of <i>A.terreus</i> IBT15722 gDNA spanning putative ORFs ATEG_8453-8454 including 1788 bp overlap with 455/456 PCR product
9	84/71	atctccaguTTACACA GACCAGCAATCAT	agtgagcuCGCGGTTTT TTGGGGTAGTCATC	<i>A. nidulans</i> cloning vector pU1111 (Hansen et al, 2011)
10	77/422	atctccaguAGCGCG GCCGCATTAAATC	acagagtucuTGAAGTG GTGGGCTAACTAC	1,6 kb of basic cloning vector PJ204 (DNA 2.0) including the <i>ampR</i> marker gene

11	421/70	agaactctguAGCACC GCCTACATACCTCG	agtgagcuTGGATAAC CGTATTACCGCCTTTG	5,3 kb of <i>A. nidulans</i> cloning vector pU2000 (nomenclature from (20)) including <i>A. fumigatus pyrG</i> marker gene flanked by direct repeats and <i>IS1</i> downstream targeting region
12	BGHA1 63/568	GGTCTACTCCCCTG CGTCTA	CTCATCCAGCCTGTTA TCCA	Analytical PCR for verification of insertion of fragment1 in <i>IS1</i> upstream targeting site
13	BGHA9 8/162	GGTTTCGTTGTCAA TAAGGGAA	GTTTCAGGGTGACGGT GAGAG	Analytical PCR for verification of insertion of fragment1 in <i>IS1</i> downstream targeting site
14	112/10 8	gggtttaauATGTCTG CAGAAGGACACAA GC	gggtttaauATGACTAC GACCTATGCAGT	Analytical PCR for verification of insertion of fragment2 in the overlap with the fragment1
15	BGHA 182/16 2	CACACATCGACATC CTCACC	GTTTCAGGGTGACGGT GAGAG	Analytical PCR for verification of insertion of fragment2 in <i>IS1</i> downstream targeting site
16	569/57 0	TCTGATCTGTCTGG AGCCT	gatccccgggaattgccatg CCGGACTAGACGGTT GGATT	Deletion of the geodin polyketide synthase, ATEG_8451. Amplifies 2,0 kb of <i>A. terreus</i> IBT15722 gDNA upstream of the putativeORF. Specific tail for use of <i>A. fumigatus pyrG</i> as selective marker
17	571/57 2	aattccagctgaccacca tgACTGGTATCAGAT TGGTCCG	CTTCAGGTGCGAGTA ATCTT	As above, but 2,0 kb downstream of the putative ORF
18	573/57 4	CACGGACGAGCAT AGAATAG	gatccccgggaattgccatg ACAGTCAGTCTTGGT CTGCG	Deletion of the putative sulochrin halogenase, ATEG_8460. Amplifies 2,0 kb of <i>A. terreus</i> IBT15722 gDNA upstream of the putativeORF. Specific tail for use of <i>A. fumigatus pyrG</i> as selective marker
19	J58/57 6	aattccagctgaccaccatg TTACACAGACCAGCA ATCATC	AGAAGAGGTCTGTGA TGCTGA	As above, but 2,0 kb downstream of the putative ORF
20	JBN 5A/2K	catggcaattcccgggga tcTGGATAACCGTAT TACCGCC	GGAAGAGAGGTTCA CACC	Amplification of 1455 bp <i>A. fumigatus PyrG</i> 5' region for fusion PCR deletion substrates.
21	JBN 4Q/5B	TGATACAGGTCTCG GTCCC	catggtggtcagctggaatt tGCCAAGCTTAACGC GTACC	Amplification of 1217 bp <i>A. fumigatus PyrG</i> 3' region for fusion PCR deletion substrates.
22	327/32 8	GAATCACTTGCCAT CTTGCT	gatccccgggaattgccatg ATTCGGACATAGCT CGTTG	Deletion of <i>A. nidulans mdp</i> gene cluster (<i>mdpA-L</i>) - upstream fragment
23	329/33 0	aattccagctgaccacca tgTGCTGTATATGG GTCTTGG	CTGTTTGTAGGCGGT TCC	Deletion of <i>A. nidulans mdp</i> gene cluster (<i>mdpA-L</i>) - downstream fragment
24	BGHA 192/23	TCGTC AAGGTTTTTC CCTTTGT	ACGCTATTGATGTGG TGAAT	Analytical PCR for verification of <i>mdpA-L</i> deletion

Appendix G

Supporting information for Chapter 12

Over-expression of Transcription Factor in *Aspergillus aculeatus* Activates Biosynthetic Pathway and Reveals Novel Compounds Acurin A and B.

Lene M. Petersen, Dorte K. Holm, Charlotte H. Gotfredsen, Uffe H. Mortensen, and Thomas O.

Larsen.

SUPPORTING INFORMATION

Figures:

Fig. S1. Mass spectra of compounds (2), (3), (5), (6), and (7), cultivated on YES or YES + ^{13}C -6-MSA.

Fig. S2. Effect of increasing concentrations of acurin A on CLL cell viability in a CellTiter-Glo® assay. The relative cell viability is compared to DMSO control (1 %).

Tables:

Table S1. NMR Data for acurin A

Table S2. NMR Data for acurin B

Figure S1. Mass spectra of compounds (2), (3), (4), (5), and (6), cultivated on YES or YES + ^{13}C -6-MSA.

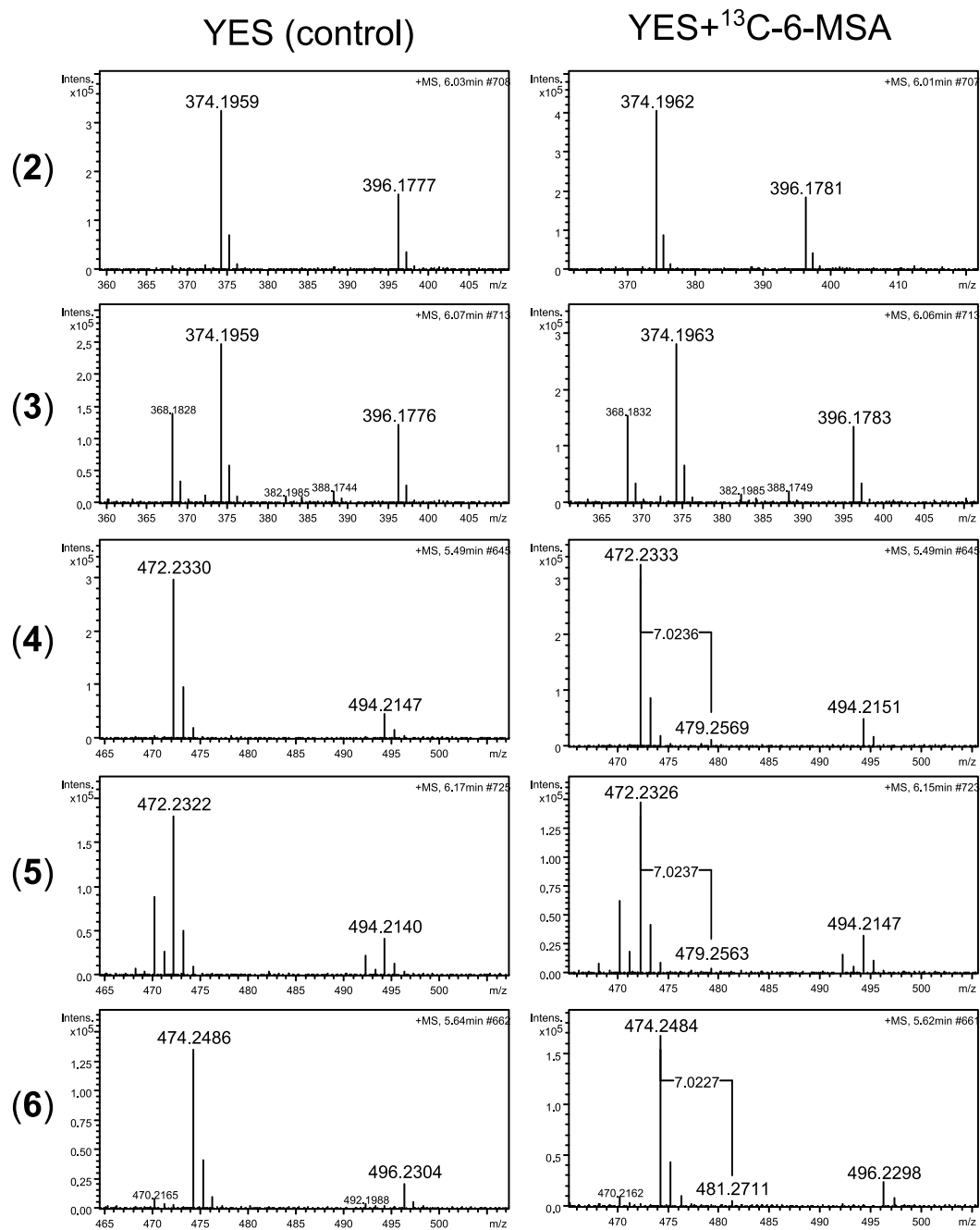


Figure S2. Effect of increasing concentrations of acurin A on CLL cell viability in a CellTiter-Glo® assay. The relative cell viability is compared to DMSO control (1 %).

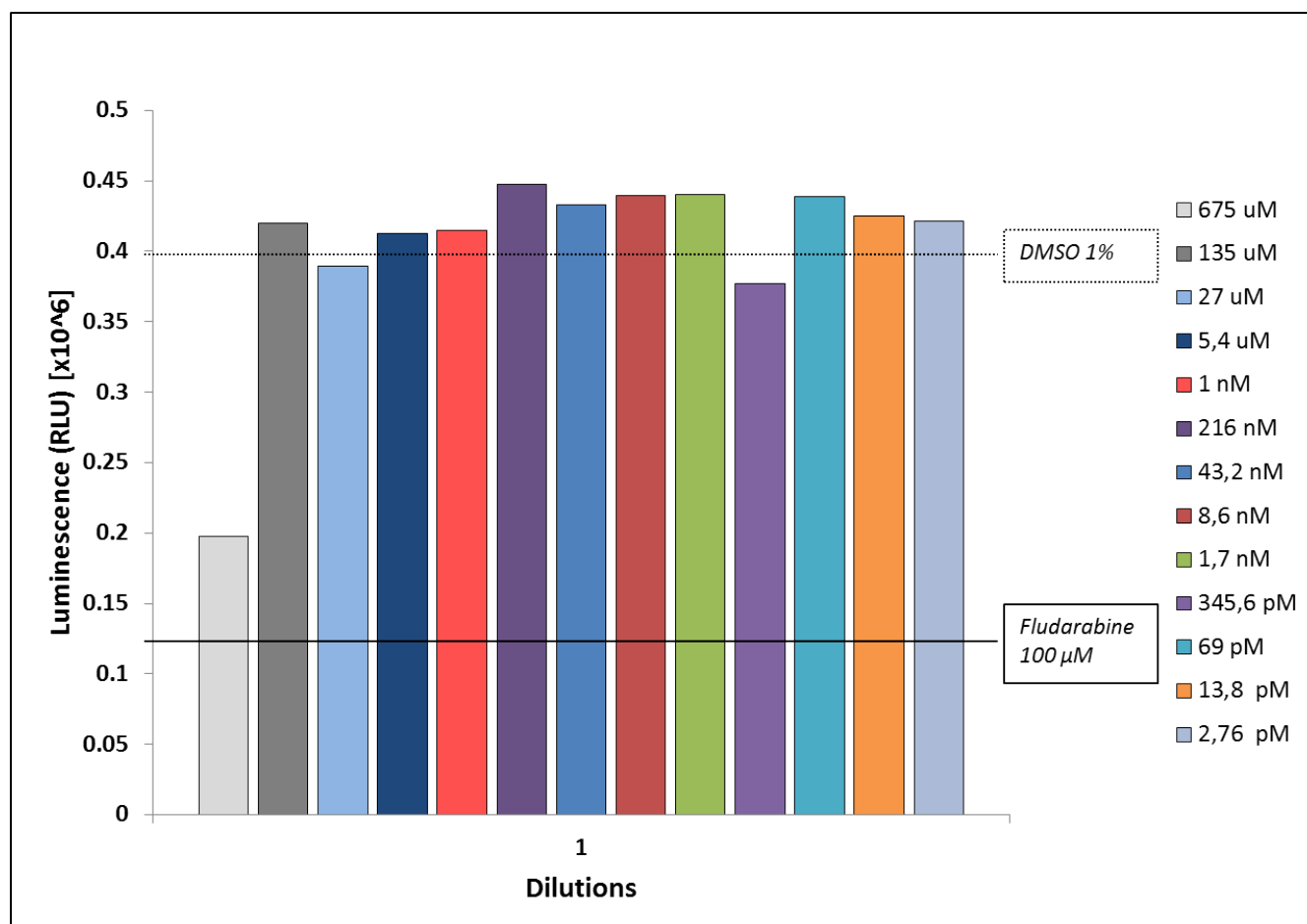
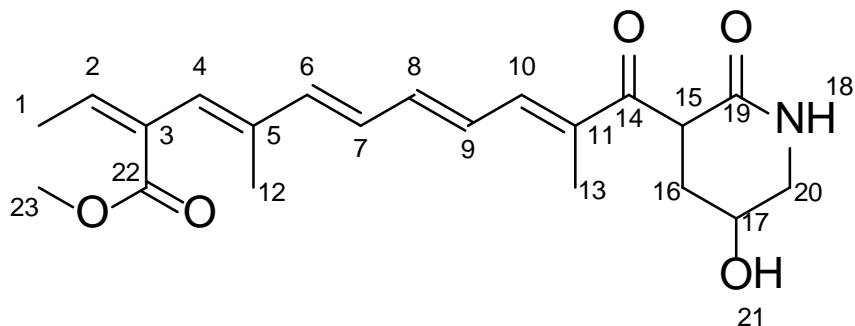
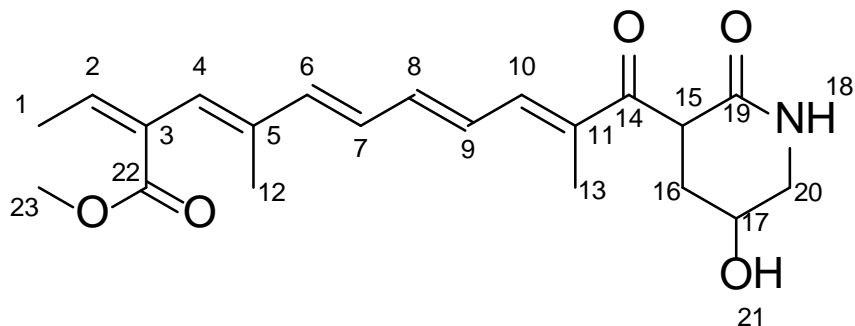


Table S1. NMR Data for acurin A in DMSO- d_6 . ^1H NMR spectrum and all 2D spectra were recorded at the Varian Unity Inova-500 spectrometer at DTU. Signals were referenced by solvent signals for DMSO- d_6 at $\delta_{\text{H}} = 2.49$ ppm and $\delta_{\text{C}} = 39.5$ ppm.



Atom	^1H -chemical shift [ppm]/ J coupling constants [Hz]	^{13}C -chemical shift [ppm]	HMBC correlations	H2BC correlations	NOE connectivities
1	1.71 (dd 7.2, 1.0)	15.4	2, 3, 22	2	2, 4
2	6.88 (m)	139.9	1, 4, 22	1	1
3	-	129.8	-	-	-
4	6.22 (s)	126.8	6, 12, 22	-	1, 6, 12
5	-	137.5	-	-	-
6	6.69 (d 15.4)	140.0	4, 5, 8, 9, 12	7	4, 7
7	6.60 (m)	128.8	5, 8	6	6, 12
8	6.83	141.0	10	7/9	6, 7
9	6.831 (m)	128.9	7	8	-
10	7.38 (m)	141.3	8, 13, 14	9	9, 15
11	-	134.4	-	-	-
12	1.64 (s)	13.8	4, 5	4	4, 7
13	1.85 (s)	11.6	10, 11, 14	10, 11	-
14	-	197.8	-	-	-
15	4.39 (m)	47.5	14, 16/16', 19	16/16'	16'
16	2.33 (m)	27.3	14, 15, 17, 20/20'	15, 20/20'	16', 20
16'	1.92 (m)	27.3	-	15, 20/20'	15, 16, 20
17	3.35 (t 5.5)	64.4	16/16'	20	-
18	7.84 (d 10)	-	15, 16/16', 19, 20/20'	-	20, 20'
19	-	173.0	-	-	-
20	3.54 (m)	53.6	16/16'	16/16', 17	16/16', 18
20'	3.36 (m)	53.6	-	-	18
21	-	-	-	-	-
22	-	166.3	-	-	-
23	3.66 (s)	51.6	22	-	-

Table S2. NMR Data for acurin B in DMSO- d_6 . ^1H NMR spectrum and all 2D spectra were recorded at the Varian Unity Inova-500 spectrometer at DTU. Signals were referenced by solvent signals for DMSO- d_6 at $\delta_{\text{H}} = 2.49$ ppm and $\delta_{\text{C}} = 39.5$ ppm.



Atom	^1H -chemical shift [ppm]/ J coupling constants [Hz]	^{13}C -chemical shift [ppm]	HMBC correlations
1	1.71 (1H, m)	14.3	2, 3, 4, 22
2	6.87 (1H, m)	139.2	1
3	-	125.0	-
4	6.84 (1H, m)	128.5	22
5	-	136.7	-
6	6.69 (dd, 15.1, 2.5)	140.0	5, 12
7	6.57 (1H, m)	137.9	3
8	7.04 (1H, m)	123.3	-
9	6.47 (1H, m)	123.7	8, 14
10	7.83 (1H, m)	135.1	13
11	-	140.0	-
12	1.63 (3H, m)	13.2	3, 5
13	1.85 (3H, s)	10.8	10, 11, 14
14	-	199.8	-
15	4.50 (1H, m)	47.3	14, 16/16', 19
16	2.35	26.6	14
16'	1.92	26.6	14
17	3.55 (2H, m)	52.9	15, 16/16', 17
18	7.89 (1H, m)	-	-
19	-	172.2	-
20	3.36 (2H, m)	63.5	-
21	-	-	-
22	-	166.3	-
23	3.65 (3H, s)	50.9	22

Appendix H

Investigation of the secondary metabolite potential in *Aspergillus aculeatus*

This chapter describes a basic bioinformatic study to reveal the secondary metabolite synthases that are present in *A. aculeatus*. The genome of *A. aculeatus* ATCC16872 was recently sequenced by the DOE Joint Genome Institute, and the automatically annotated genome is available at <http://genome.jgi-psf.org/Aspac1/Aspac1.home.html>.

H.1 Identification of synthases

For identification of synthases, I used a combination of BLAST search using characterized synthases from *A. nidulans* and keyword search for specific domains of the three types of synthases: PKSs, NRPSs, and terpenoid synthases. The search identified 24 PKSs, 19 NRPSs, 9 terpenoid synthases, and 5 hybrid synthases (H.1).

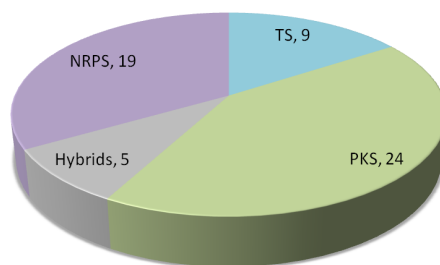


Figure H.1: Distribution of secondary metabolite synthases in *A. aculeatus*.

For identification of functional domains, the 57 putative synthases were analyzed using the Conserved Domain Database (CDD) [Marchler-Bauer et al., 2011]. The deduced domain organization is shown for all synthases in Table H.1. It should be noted that some recently discovered and poorly characterized domains

are not detected by CDD e.g. product template (PT) and thio Hydrolase (TH) domains, why these have to be annotated manually.

The putative NRPSs were additionally analyzed using the web-based program NRSPredictor2 [Röttig et al., 2011] which allows for prediction of the adenylation domain specificity i.e. which amino acid is selected and attached to the NRP chain.

Table H.1: Secondary metabolite synthases in *A. aculeatus*. ^aPredicted by AUGUSTUS, ^bPredicted by FGeneSH, ^cPredicted amino acid recognized by the adenylation domain (Score: 0-100) by NRPS Predictor2 [Röttig et al., 2011]. TIDS = trans-isoprenyl diphosphate synthase, OHFT = Omega-hydroxypalmitate O-feruloyl transferase, Dhb = 2,3-dihydroxy-benzoic acid, Hpg = 4-hydroxy-phenyl-glycine, Orn = ornithine, Aad = 2-amino-adipic-acid.

Type	Model ID	Length (aa)	Domains (CDD/NRPS2)
PKS	48	2429	KS AT DH MT DH KR ACP
PKS	108	2409	KS AT DH MT DH ACP
PKS	26431	2426	KS AT MT DH KR ACP
PKS	28570	2548	KS AT MT DH KR ACP
PKS	29093	2387	KS AT MT DH KR ACP
NR-PKS	31261	2165	KS AT ACP TE
NR-PKS	33832	1793	KS AT ACP/thiolase-like
PKS	42038	2896	KS AT DH MT DH KR ACP DH
PKS	42746	2186	KS AT DH MT KR ACP
PKS	43562	2372	KS AT DH ER KR ACP
PKS	63275-46489	213	KR ACP (ER KR ACP) ^a
PKS	47290	2422	KS AT DH MT ER KR ACP
PKS	48557	3214	KS AT DH MT ER DH ACP Monoamine oxidase
PKS	50671	2346	KS AT DH ER KR
NR-PKS	54632	2546	KS AT ACP MT TR
PKS	55425	2437	KS AT DH MT ER KR
PKS	63311	3042	PKS - DH ACP DH
PKS	65031	1681	AT ACP KR C
PKS	120020	429	KS
NR-PKS	126899	2143	SAT KS AT ACP TE
PKS	1867870	2470	KS AT DH MT KR ACP
PKS	1870238	2467	KS AT DH MT ER KR ACP
NR-PKS	1874409	1779	KS AT ACP

Continued on Next Page

Table H.1 – Continued

Type	Model ID	Length (aa)	Domains (CDD/NRPS2)
PKS	1882131(*)	563 (2247)	DH KR ACP (OHFT MT ER KR ACP) ^b
PKS	1889732	2165	IPP isomerase - KS AT MT DH - TIDS
PKS	1904397	1750	KS AT DH KR ACP
NRPS	20389	1022	Cys (50) ^c
NRPS	19967	4728	Ala (90), Dhb (90), ser (90)
NRPS	241	1780	Leu (40), Arg (40)
NRPS	1878108	1734	Val (70), Aad (60)
NRPS	112572	1427	Pro (90)
NRPS	802	1066	Val (50)
NRPS	115976	1998	Cys (60)
NRPS	42025	7330	Hpg (40), Ala (60), Dhb (40), Val (50), Ala (60), Tyr (40)
NRPS	60230	6264	Pro (60), Ser (40), Orn (60), Val (70)
NRPS	43289	1071	Leu (50)
NRPS	29371	1079	Leu (60)
NRPS	122295	1553	Hpg (50)
NRPS	34099	2140	Lys (30), Leu (40)
NRPS	394	1477	Asp (40), Cys (60)
NRPS	64101	2407	Thr (40), Ala (60)
NRPS	64146	5041	Phe (60), Cys (50), Orn (50), Gln (40)
NRPS	1907395	2068	Val (60)
NRPS	1860611	5438	Phe (40), Phe (50), Pro (50), Pro (60), Leu (50)
NRPS	37457	7009	Cys (60), Leu (70), Val (70), Pro (40), Glu (50)
PKS/NRPS	51108	4073	KS AT DH MT KR ACP C A P DH (Glu (50))
PKS/NRPS	43247	3970	KS AT DH KR ACP C A P DH (Val (50))
PKS/NRPS	53248	4018	KS AT DH KR ACP C A P DH (Ser(100))
PKS/NRPS	31401	3937	KS AT MT DH KR ACP C A P DH (Val (100))
Terpene synthase	48138	210	Terpene cyclase
Terpene synthase	57499	269	Terpene cyclase
Terpene synthase	58369	365	
Terpene synthase	58711	339	Terpene cyclase
Terpene synthase	60731	443	Terpene cyclase
Terpene synthase	116070	373	Terpene cyclase
Terpene synthase	125390	385	Terpene cyclase

Continued on Next Page

Table H.1 – Continued

Type	Model ID	Length (aa)	Domains (CDD/NRPS2)
Terpene synthase	127866	744	Polyprenyl synthetase
Terpene synthase	1858621	778	Polyprenyl synthetase

H.2 Annotation and assessment of polyketide synthases

The *A. aculeatus* genome contains 24 putative PKSs, of which only five appear to be non-reducing (NR) i.e. they do not contain any reduction domains. A very interesting discovery was the presence of a hybrid IPP isomerase-PKS-trans-isoprenyl diphosphate synthase (1889732). An IPP isomerase is the enzyme that catalyzes the reversible reaction of isopentenylpyrophosphate (IPP) into dimethylallylpyrophosphate (DMAPP) in the biosynthesis towards terpenoids. Trans-isoprenyl diphosphate synthases (TIDSs) catalyze head-to-tail condensation of IPP and DMAPP building blocks to synthesize terpenes. To my knowledge, this configuration has not been reported before, and although it is not possible to predict the product, it should be fair to assume that it is a hybrid meroterpenoid (PKS-terpene) which is synthesized in a novel manner than previously reported meroterpenoids e.g. austinols.

Although most of the PKS seem to be full-length (containing a full set of domains), three putative synthases are very short and do not contain a full set of domains: 46489, 120020, and 1882131.

H.2.1 46489 is a non-ordinary or non-functional PKS

CDD analysis suggested that 46489 comprised two domains: KR (partial) and ACP (Figure H.3A). As the ACP is normally found in the C-terminus of the PKS, I examined the upstream genome sequence for PKS activities. One gene (63275), located just upstream of 46489, did indeed encode an enoyl reductase (ER) (Figure H.3B) which is present in highly reducing PKSs. The entire genomic stretch surrounding 46489 (+/- 20 kb) was analyzed using the two gene prediction algorithms FgeneSH (Softberry) and AUGUSTUS [Stanke et al., 2004], and AUGUSTUS did indeed predict the two genes as one single gene, encoding three full domains: ER, KR, and ACP (Figure H.3C). However, I was unable to locate any KS or AT domains, which are essential for ordinary PK synthesis by a type I iterative PKS. Thus this putative PKS must be a non-ordinary PKS, or a non-functional part of a type I iterative PKS which has been partly lost or partly gained by an unknown mechanism.

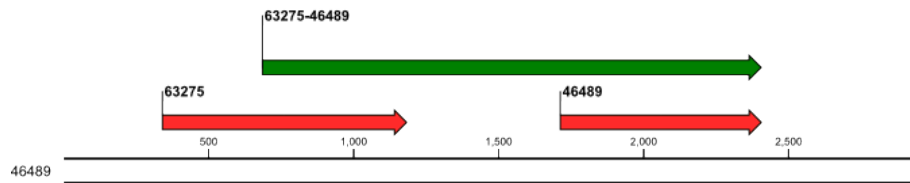


Figure H.2: Gene prediction for putative PKS 46489. JGI prediction in red, AUGUSTUS prediction in green.



Figure H.3: Annotation of putative PKS 46489. **A.** CDD search for 46489 identifies KR and ACP domains. **b.** CDD search for 63275 identifies ER domain. **C.** CDD search for combined 63275-46489 (predicted by AUGUSTUS) identifies ER, KR and ACP domains.

H.2.2 120020 may be a type III PKS

The putative PKS 120020 was analysed by CDD, revealing that only a KS domain was present. Analysis of the genome sequence surrounding the putative PKS gene did not reveal any additional domains, however several non-PKS-related genes were predicted. This suggested that the PKS did in fact only comprise a KS domain, which led me to propose that this gene may encode a type III PKS.

In contrast to the multidomain type I and type II PKSs, type III PKSs are selfcontained homodimeric enzymes that on their own can catalyze a whole set of reactions: recruitment of a starter unit, decarboxylative condensation of extender units, ring closures, and/or cyclization of the polyketide chain [Austin and Noel, 2009]. All of these reactions take place within the same active site pocket.

Type III PKSs are most common in bacteria but interestingly, Li and co-workers recently identified and characterized a type III PKS from *A. niger* [Li et al., 2011a]. As this type of PKSs is in fact present in *Aspergilli*, it is likely that 120020 is a type III PKS.

H.2.3 1882131 is a non-ordinary PKS that contains a novel SAT domain

The 1882131 putative PKS was analyzed by CDD, and only three domains were identified: DH, KR, and ACP (Figure H.5A). Analysis of the surrounding genome sequence (+/- 20 kb) and subsequent re-annotation by FGeneSH and AUGUSTUS revealed that the ORF was in fact longer than initially predicted (Figure H.5B). However, no KS nor AT domain was identified. Instead, the predicted gene contained part of a conserved domain from a omega-hydroxypalmitate O-feruloyl transferase (OHFT) in the N-terminus. A OHFT catalyzes the transfer of a feruloyl-CoA to a 16-hydroxypalmitate, yielding a 16-feruloyloxypalmitate (Figure H.6) [Gou et al., 2009, Molina et al., 2009].



Figure H.4: Gene prediction for putative PKS 1882131. JGI prediction in red, AUGUSTUS and FGeneSH prediction in green.

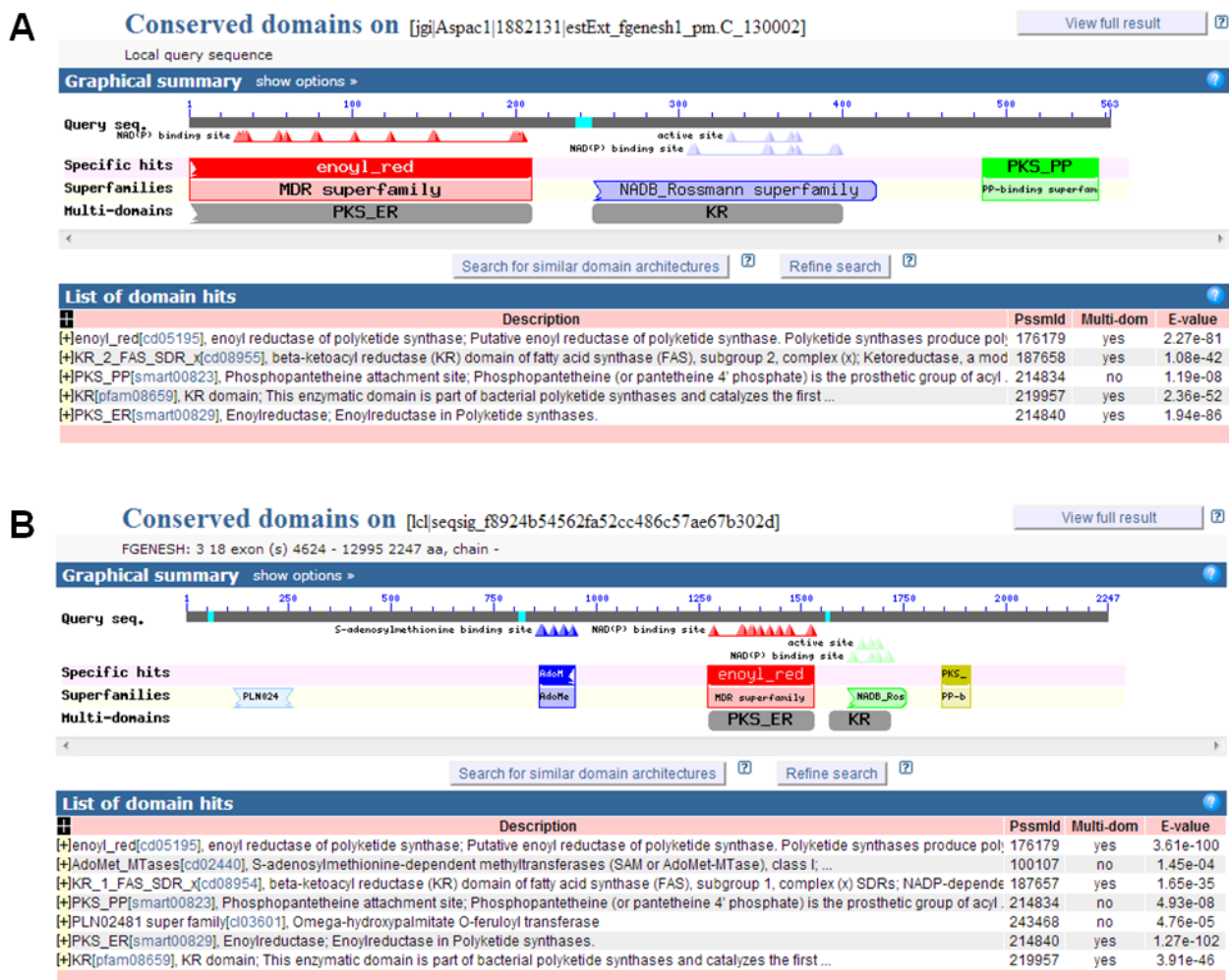


Figure H.5: Annotation of putative PKS 1882131. **A.** CDD search for 1882131 identifies ER, KR and ACP domains. **B.** CDD search for 1882131* (predicted by FgeneSH) identifies OHFT, MT, ER, KR and ACP domains.

Interestingly, there is a non-identified stretch of >500 amino acids between the predicted OHFT and the methyl transferase (MT) active site. This stretch could contain novel domains which obviate the need for KS and AT domains, which are normally present in a functional PKS. Although I cannot speculate concerning the products of the PKS, I propose that 1882131 is a functional, non-ordinary PKS that employs an alternative starter unit which is recruited and attached by the alternative transferase domain.

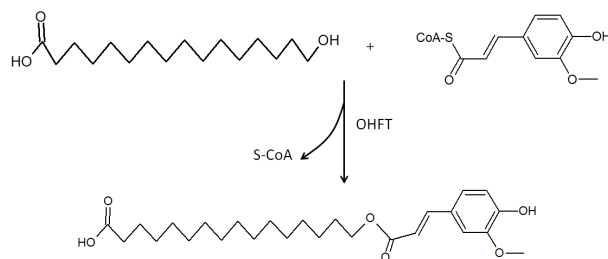


Figure H.6: Activity of the omega-hydroxypalmitate O-feruloyl transferase (OHFT). OHFT attaches the feruloyl-CoA to the C16-hydroxy group of 16-hydroxypalmitate to yield 16-feruloyloxypalmitate.

H.3 Phylogeny of the *A. aculeatus* PKSs

in order to get an overview of the PKSs that are present in *A. aculeatus*, I aligned the amino acid sequences using ClustalW and made a phylogenetic tree (Figure H.7) [Larkin et al., 2007]. For reference, I included *A. nidulans wA* (AN8209) and *A. niger yanA* (ASPNI44965). As expected, the reducing PKSs cluster together, and the non-reducing PKSs cluster together. There are four PKSs that do not cluster with the two large groups: 65031, 120020, 1999732, and 1904397. The latter clusters with *yanA*, and further analyses have verified that 1904397 is in fact a 6-MSA synthase (Chapter 12). 65031 and 120020 do not seem complete, which could explain the lack of homology to the remaining PKSs. 1889732 is described in Section H.2 as a hybrid PKS-terpene synthase, and as this is a novel configuration it could certainly explain why it differs from the other PKSs.

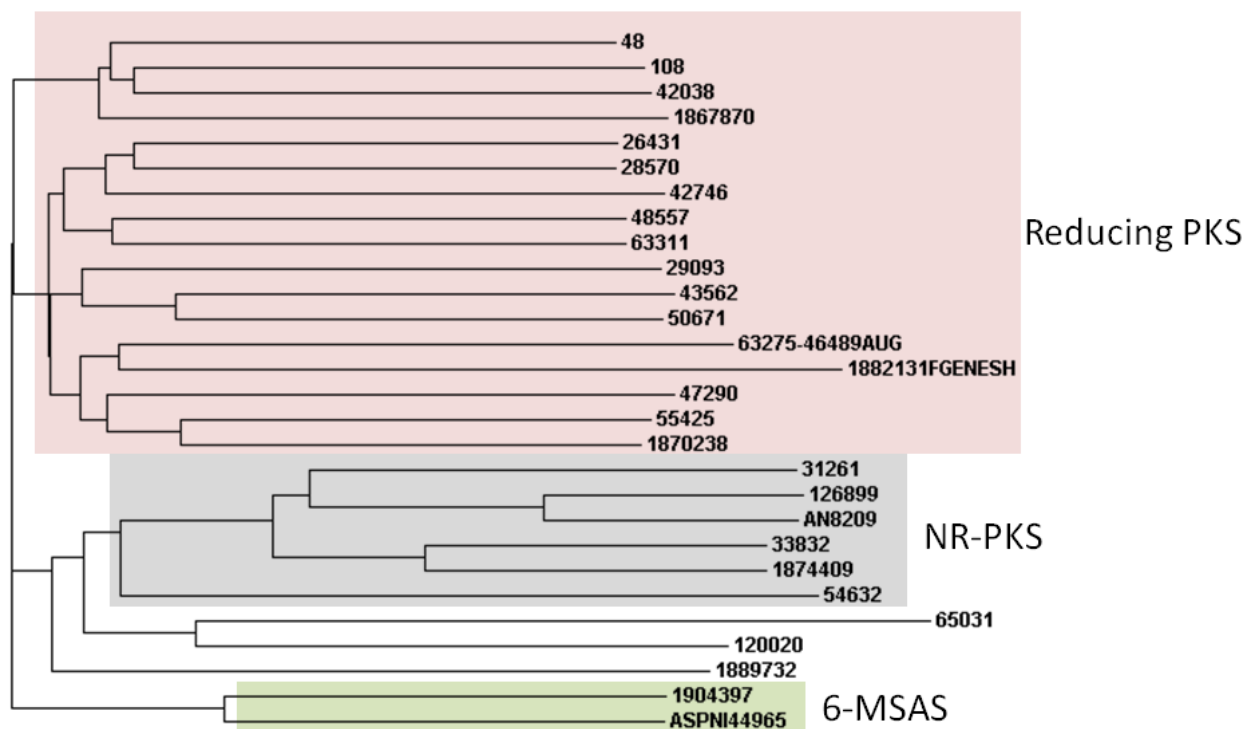


Figure H.7: Phylogenetic tree of *A. aculeatus* PKSs with two references: ASPNI44965 (*A. niger* 6-MSA synthase *yanA*) and AN8209 (*A. nidulans* YWA1 synthase *wA*).